

Министерство образования и науки Российской Федерации
федеральное государственное автономное образовательное учреждение
высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Инженерная школа природных ресурсов
Направление подготовки 21.04.01 «Нефтегазовое дело»
Отделение нефтегазового дела

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема работы
Типизация геолого-геофизических разрезов на основе анализа каротажных данных с использованием искусственных нейронных сетей на примере месторождений Нюрольской впадины

УДК 550.832:004.7.032.26 (571.1)

Студент

Группа	ФИО	Подпись	Дата
2ТМ61	Канаев Илья Сергеевич		

Руководитель ВКР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
инженер ЦППС НД ТПУ	Коровин М.О.	К.Г.-М.Н.		

КОНСУЛЬТАНТЫ

По разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
профессор	Белозеров В.Б.	Д.Г.-М.Н.		

По разделу «Социальная ответственность»

Должность	ФИО	Ученая степень, звание	Подпись	Дата
ведущий эксперт ЦППС НД ТПУ	Меркулов В.П.	К.Г.-М.Н.		

ДОПУСТИТЬ К ЗАЩИТЕ

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
ведущий эксперт ЦППС НД ТПУ	Чернова О.С.	К.Г.-М.Н.		

Министерство образования и науки Российской Федерации
 Федеральное государственное автономное образовательное учреждение
 высшего образования
**«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
 ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

Инженерная школа природных ресурсов
 Направление подготовки 21.04.01 «Нефтегазовое дело»
 Отделение нефтегазового дела

УТВЕРЖДАЮ:
 Руководитель ООП

 (Подпись) (Дата) Чернова О.С.
 (Ф.И.О.)

ЗАДАНИЕ
на выполнение выпускной квалификационной работы

В форме:

Магистерской диссертации

Студенту:

Группа	ФИО
2ТМ61	Канаев Илья Сергеевич

Тема работы:

Типизация геолого-геофизических разрезов на основе анализа каротажных данных с использованием искусственных нейронных сетей (на примере месторождений Нюрольской впадины)	
Утверждена приказом директора (дата, номер)	08.08.2018 №7260/с

Срок сдачи студентом выполненной работы:	18.08.2018
--	------------

ТЕХНИЧЕСКОЕ ЗАДАНИЕ:

Исходные данные к работе	1. Данные геофизических исследований скважин с исследуемых месторождений Нюрольской впадины 2. Набор геологической информации по скважинам 3. Результаты исследований кернового материала 4. Набор данных о проведенных промысловых исследованиях 5. Данные полученные при бурении скважин
Перечень подлежащих исследованию, проектированию и разработке вопросов	1. Анализ применимости методов машинного обучения для решения поставленных в работе задач 2. Изучение свойств и возможностей нейронных сетей 3. Анализ возможности автоматического выделения реперных горизонтов с применением методов машинного обучения 4. Изучение научных работ и исследовательского опыта в данной области 5. Препроцессинг исходных данных 6. Применение нейронных сетей для решения задачи крупномасштабной классификации комплекса горных пород 7. Реализация других методов машинного обучения для целей сравнения результатов 8. Анализ результатов полученных с применением

	нейронных сетей и других методов классификации 9. Реализация метода выделения реперных горизонтов на основе каротажных данных с применением методов машинного обучения 10. Комплексный анализ полученных результатов 11. Выполнить задание блока «Финансовый менеджмент. Ресурсоэффективность и ресурсосбережение.» 12. Рассмотреть вопросы связанные с социальной ответственностью
Перечень графического материала	Визуализация входных данных, конфигурация сети, матрицы результатов классификации, визуализация итогов, карты тектонических нарушений, временные разрезы, геологические разрезы.

Консультанты по разделам выпускной квалификационной работы

Раздел	Консультант
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	профессор Белозеров В.Б.
Социальная ответственность	ведущий эксперт ЦППС НД ТПУ Меркулов В.П.

Названия разделов, которые должны быть написаны на русском и иностранном языках:

Введение Project introduction
Анализ предметной области и актуальность выполняемой работы Object domain analysis and relevance of project
Анализ геологической изученности района исследований Geological exploration analysis of interest area
Анализ данных Data analysis
Конфигурация классификатора Configuration of the classifier
Анализ результатов исследования Results analysis
Финансовый менеджмент, ресурсоэффективность и ресурсосбережение Financial management, resource efficiently and resource conservation
Социальная ответственность Social responsibility
Заключение Conclusion

Дата выдачи задания на выполнение выпускной квалификационной работы по линейному графику	27.06.2018
---	------------

Задание выдал руководитель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
инженер ЦППС НД ТПУ	Коровин М.О.	к.г.-м.н.		

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
2ТМ61	Канаев Илья Сергеевич		

ЗАПЛАНИРОВАННЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ООП
21.04.01 «Нефтегазовое дело»

№	Результаты обучения
1	2
P1	Применять естественнонаучные, математические, гуманитарные, экономические, инженерные, технические и глубокие профессиональные знания в области современных нефтегазовых технологий для решения прикладных междисциплинарных задач и инженерных проблем, соответствующих профилю подготовки
P2	Планировать и проводить аналитические и экспериментальные исследования с использованием новейших достижений науки и техники, уметь критически оценивать результаты и делать выводы, полученные в сложных и неопределённых условиях; использовать принципы изобретательства, правовые основы в области интеллектуальной собственности
P3	Проявлять профессиональную осведомленность о передовых знаниях и открытиях в области нефтегазовых технологий с учетом передового отечественного и зарубежного опыта; использовать инновационный подход при разработке новых идей и методов проектирования объектов нефтегазового комплекса для решения инженерных задач развития нефтегазовых технологий, модернизации и усовершенствования нефтегазового производства
P4	Выбирать оптимальные решения в многофакторных ситуациях, владеть методами и средствами технического моделирования производственных процессов и объектов нефтегазовой отрасли; управлять технологическими процессами, обслуживать оборудование, использовать любой имеющийся арсенал технических средств, обеспечивать высокую эффективность при разработке нефтегазовых объектов
P5	Самостоятельно учиться и непрерывно повышать квалификацию в течение всего периода профессиональной деятельности; активно владеть иностранным языком на уровне, позволяющем работать в интернациональной среде, разрабатывать документацию и защищать результаты инженерной деятельности
P6	Работать эффективно в качестве члена и руководителя команды, умение формировать задания и оперативные планы всех видов деятельности, распределять обязанности членов команды, нести ответственность за результаты работы; координировать работу групп по извлечению и совершенствованию добычи нефти, газа и газового конденсата, передавать знания через наставничество и консультирование
P7	Внедрять, эксплуатировать и обслуживать современные машины и механизмы для реализации техно-логических процессов нефтегазовой области, обеспечивать их высокую эффективность, соблюдать правила охраны здоровья и безопасности труда, выполнять требования по защите окружающей среды
P8	Проявлять профессиональную осведомленность о передовом отечественном и зарубежном опыте в моделировании динамически вязких сред и низкопроницаемых коллекторов
P9	Предлагать процедуры оценки эффективности промысловых работ и оптимизации работы оборудования при добыче нефти, газа и газового конденсата, обеспечение энергоэффективности технологических процессов
P10	Обеспечивать внедрение новых методов, материалов и нефтегазового оборудования в осложненных условиях эксплуатации нефтяных и газовых скважин, прогнозировать режимы безопасной работы нефтегазового оборудования по динамическим, локальным и осредненным параметрам
P11	Контролировать выполнение требований регламентов для обеспечения добычи нефти, газа и газового конденсата и повышение интенсификации притока скважинной продукции
P12	Совершенствовать, разрабатывать мероприятия и/или подготавливать бизнес-предложения по технологическому процессу и технологическим мероприятиям при добыче нефти, газа и газового конденсата на основе производственного менеджмента и планирования работ в сфере нефтегазодобычи
P13	Корректировать программы работ по добыче нефти, газа и газового конденсата, выбирать и принимать решения в нестандартных ситуациях, опираясь на государственные стандарты в области нефтегазодобычи

Министерство образования и науки Российской Федерации
Федеральное государственное автономное образовательное учреждение
высшего образования
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ПОЛИТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»

Инженерная школа природных ресурсов
 Направление подготовки 21.04.01 «Нефтегазовое дело»
 Уровень образования: магистратура
 Отделение нефтегазового дела
 Период выполнения: весенний семестр 2018 учебного года
 Форма представления работы:

Магистерская диссертация

КАЛЕНДАРНЫЙ РЕЙТИНГ-ПЛАН
выполнения выпускной квалификационной работы

Срок сдачи студентом выполненной работы:	18.08.2018
--	------------

Дата контроля	Название раздела (модуля) / вид работы (исследования)	Максимальный балл раздела (модуля)
03.07.2018	Анализ предметной области и актуальность работы	10
13.07.2018	Анализ геологической изученности района исследований	15
20.07.2018	Анализ данных	20
25.07.2018	Конфигурация классификатора	20
01.08.2018	Анализ результатов исследования	15
08.08.2018	Социальная ответственность	10
13.08.2018	Финансовый менеджмент, ресурсоэффективность и ресурсосбережение	10

Составил преподаватель:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
инженер ЦППС НД ТПУ	Коровин М.О.	к.г.-м.н.		27.06.2018

СОГЛАСОВАНО:

Руководитель ООП	ФИО	Ученая степень, звание	Подпись	Дата
ведущий эксперт ЦППС НД ТПУ	Чернова О.С.	к. г.—м. н.		27.06.2018

**ЗАДАНИЕ ДЛЯ РАЗДЕЛА
«ФИНАНСОВЫЙ МЕНЕДЖМЕНТ, РЕСУРСОЭФФЕКТИВНОСТЬ И
РЕСУРСОСБЕРЕЖЕНИЕ»**

Студенту:

Группа	ФИО
2ТМ61	Канаеву Илье Сергеевичу

Школа	ИШПР	Отделение школы	нефтегазового дела
Уровень образования	магистратура	Направление/специальность	21.04.01 Нефтегазовое дело/ Разработка и эксплуатация нефтяных и газовых месторождений

Исходные данные к разделу «Финансовый менеджмент, ресурсоэффективность и ресурсосбережение»:

1. Стоимость ресурсов научного исследования (НИ): материально-технических, энергетических, финансовых, информационных и человеческих	Стоимость материальных ресурсов определялась по средней стоимости по г. Томску. Оклады в соответствии с окладами сотрудников НИ ТПУ: доцент, к.г.–м.н. – 23600 руб.
2. Нормы и нормативы расходования ресурсов	Премияльный коэффициент 15%; коэффициент доплат и надбавок 20%; коэффициент дополнительной заработной платы 15%; районный коэффициент 30%
3. Используемая система налогообложения, ставки налогов, отчислений, дисконтирования и кредитования	Коэффициент отчислений на уплату во внебюджетные фонды 31,8 %

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

1. Оценка коммерческого и инновационного потенциала НТИ	Анализ преимуществ реализации проекта
2. Разработка устава научно-технического проекта	Анализ потенциальных потребителей, сегментирование рынка
3. Планирование процесса управления НТИ: структура и график проведения, бюджет, риски и организация закупок	Определение видов работ, определение трудоемкости работ, разработка календарного плана-графика, формирование бюджета проекта, оценка рисков
4. Определение ресурсной, финансовой, экономической эффективности	Расчет интегрального показателя ресурсоэффективности

Перечень графического материала (с точным указанием обязательных чертежей):

1. «Портрет» потребителя результатов НТИ
2. Сегментирование рынка
3. Оценка конкурентоспособности технических решений
4. График проведения и бюджет НТИ
5. Реестр рисков проекта
6. Оценка ресурсной, финансовой и экономической эффективности НТИ

Дата выдачи задания для раздела по линейному графику	27.06.2018
---	------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
Профессор	Белозеров В.Б.	д.г.—м.н.		27.06.2018

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
2ТМ61	Канаев Илья Сергеевич		27.06.2018

ЗАДАНИЕ ДЛЯ РАЗДЕЛА «СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ»

Студенту:

Группа	ФИО
2ТМ61	Канаеву Илье Сергеевичу

Школа	ИШПР	Отделение	Нефтегазовое дело
Уровень образования	Магистратура	Направление/специальность	21.04.01 Нефтегазовое дело/ Разработка и эксплуатация нефтяных и газовых месторождений

Исходные данные к разделу «Социальная ответственность»:

1. Характеристика объекта исследования и области его применения	Объектом исследования является программное обеспечение для типизации геолого-геофизических разрезов с применением искусственных нейронных сетей. Произвести классификацию массива горных пород и выделить маркирующие горизонты, для произведения дальнейшей корреляции разреза. Данное исследование актуально для нефтегазовых компаний и сервисных компаний. Рабочее место должно быть оборудовано стулом, столом и расположенном на нем ПК.
---	--

Перечень вопросов, подлежащих исследованию, проектированию и разработке:

<p>Производственная безопасность</p> <p>1.1 Анализ выявленных вредных факторов при разработке и эксплуатации проектируемого решения в следующей последовательности:</p> <ul style="list-style-type: none"> – физико-химическая природа вредности, её связь с разрабатываемой темой; – действие фактора на организм человека; – приведение допустимых норм с необходимой размерностью. <p>1.2 Анализ выявленных опасных факторов при разработке и эксплуатации проектируемого решения в следующей последовательности:</p> <ul style="list-style-type: none"> – механические опасности; – термические опасности; – электробезопасность; – пожаровзрывобезопасность. 	<p>1.1 Вредные производственные факторы, создаваемые объектом исследования в процессе разработки и реализации, идентичны и возникают из-за использования ПК:</p> <ul style="list-style-type: none"> – недостаточное освещение; – нарушение параметров микроклимата; – ионизирующие излучения; – умственная перенапряженность; – монотонность работы. <p>1.2 Опасные производственные факторы, создаваемые объектом исследования:</p> <ul style="list-style-type: none"> – поражение электрическим током; – статическое электричество; – короткое замыкание.
<p>2 Экологическая безопасность</p> <p>2.1 Анализ влияния объекта исследования на окружающую среду.</p>	<p>2.1 Программное обеспечение не оказывает прямого влияния на окружающую среду, однако разрабатывается и эксплуатируется на ПК, влияющем на нее при утилизации комплектующих,</p>

	периферийного оборудования, бумажных отходов и люминесцентных ламп.
3 Безопасность в чрезвычайных ситуациях 3.1 Анализ вероятных ЧС, которые может инициировать объект исследований и которые могут возникнуть на рабочем месте при проведении исследований. 3.2 Разработка действий в результате возникшей ЧС и мер по ликвидации её последствий.	3.1 ЧС, возможные при разработке и работе программного обеспечения, зависят от персонального компьютера. Наиболее вероятные ЧС, возникающие на рабочем месте: пожары и взрывы. 3.2 Разработаны мероприятия по предотвращению наиболее типичной ЧС – пожара, согласно нормативным документам: <ul style="list-style-type: none"> – НПБ 105-03. Определение категорий помещений, зданий и наружных установок по взрывопожарной и пожарной опасности; – ППБ 01-03.
4 Правовые и организационные вопросы обеспечения безопасности 4.1 Специальные (характерные для проектируемой рабочей зоны) правовые нормы трудового законодательства. 4.2 Организационные мероприятия при компоновке рабочей зоны.	4.1 Описание правовых норм для работ, связанных с работой за персональным компьютером, согласно <ul style="list-style-type: none"> – трудовому кодексу Российской Федерации от 30.12.2001 N 197-ФЗ (ред. от 30.12.2015); – ГОСТ Р ИСО 9241-1-2007; – ГОСТ Р ИСО 9241-2-2009. 4.2 Разработаны следующие мероприятия при компоновке рабочей зоны: <ul style="list-style-type: none"> – оснащение рабочего места; – обучение рабочего персонала.

Дата выдачи задания для раздела по линейному графику	27.06.2018
---	------------

Задание выдал консультант:

Должность	ФИО	Ученая степень, звание	Подпись	Дата
профессор	Белозеров В.Б.	д.г.—м.н.		27.06.2018

Задание принял к исполнению студент:

Группа	ФИО	Подпись	Дата
2ТМ61	Канаев Илья Сергеевич		27.06.2018

РЕФЕРАТ

Выпускная магистерская диссертация содержит 100 страниц, 32 рисунка, 13 таблиц, 34 источника.

Ключевые слова: машинное обучение, нейронные сети, типизация разреза, временной ряд, скрытые Марковские модели, каскад сетей.

Целью данной магистерской диссертации является применение искусственной нейронной сети для типизации геолого-геофизического разреза. Исходными материалами для анализа являлся набор каротажных данных с месторождений, территориально приуроченных к Нюрольской впадине.

Объектами исследования являются месторождения, приуроченные к Нюрольской впадине, Западно-Сибирской нефтегазоносной мегапровинции.

В магистерской диссертации исследованы различные методы применения нейронных сетей для целей типизации геолого-геофизического разреза, рассмотрена конфигурация классификатора. Одной из важнейших частей исследования является произведение процесса идентификации маркирующих горизонтов, для целей дальнейшей корреляции геолого-геофизических разрезов скважин. В ходе выполнения магистерской диссертации было создано программное обеспечение, позволяющее производить типизацию геолого-геофизических разрезов на основании каротажных данных. Произведен анализ финансовой и ресурсной эффективности проекта реализации. Особое внимание уделено вопросам обеспечения безопасности и экологической ответственности.

Область применения: результаты, полученные в ходе выполнения данной магистерской диссертации, могут быть использованы для построения корреляционных панелей по данным месторождениям Нюрольской впадины, а также для обработки данных поступающих с введенных в эксплуатацию скважин.

Определения, обозначения и сокращения

Определения

В данной работе применены следующие термины с соответствующими определениями:

каротажные данные: Тип данных получаемых вследствие геофизических изысканий. Являются различными физическими свойствами горных пород, состояния ствола скважины или оборудования.

свита: Единица местной стратиграфической шкалы. Комплекс пород, характеризующийся определенными особенностями, образовавшийся в пределах какого-либо региона в определенных условиях и занимающий определенное стратиграфическое положение в данном регионе.

временной разрез: Способ наглядной визуализации сейморазведочных данных на плоскости в координатах: расстояние вдоль профиля - время отражения по нормали.

фораминиферы: Класс раковинных одноклеточных организмов из группы протистов. Преимущественно морские протисты; некоторые формы обитают в солоноватых и пресных водоемах

фоссилии: Ископаемые остатки организмов или следы их жизнедеятельности, относящиеся к прежним геологическим эпохам. Фоссилии предоставляют важную информацию об организмах эпохи своего образования. Фоссилии обычно представляют собой остатки или отпечатки животных и растений, сохранившиеся в массиве горных пород

тектонический разлом: Зона нарушения сплошности массива горных пород, слагающих земную кору, деформационный шов, разделяющий породный массив на два блока.

искусственные нейронные сети: Математическая модель, а также её программное или аппаратное воплощение, построенная по принципу организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма.

задача классификации: Задача отнесения образца к одному из нескольких попарно не пересекающихся множеств.

функция активации: Функция, вычисляющая выходной сигнал искусственного нейрона. В качестве аргумента принимает сигнал, получаемый на выходе входного сумматора.

формальный нейрон: Процессорный элемент, преобразователь данных, получающий входные данные и преобразующий их в соответствии с заданной функцией и параметрами.

Обозначения и сокращения

В данной работе применены следующие обозначения и сокращения:

ГК– гамма каротаж;

БК– боковой каротаж;

НГК–нейтронный гамма каротаж;

ГГК-П– гамма–гамма каротаж плотностной;

ИК– индукционный каротаж;

ПС– каротаж потенциалов собственной поляризации;

АК– акустический каротаж;

ДС–диаметр ствола скважины, калипер;

ЧС – чрезвычайная ситуация;

ПБ – пожарная безопасность;

ТПУ – Томский политехнический университет;

ПФРФ – пенсионный фонд Российской Федерации;

ФЗП– фонд заработной платы;

ФСС – фонд социального страхования;

ФФОМС – федеральный фонд обязательного медицинского страхования;

ГОСТ– государственный общесоюзный стандарт;

СанПиН– санитарно-эпидемиологические правила и нормы.

Введение.....	15
1 Анализ предметной области и актуальность выполняемой работы	17
2 Анализ геологической изученности района исследований	23
2.1 Географическое расположение и изученность территории	23
2.2 Геологическое строение территории исследования.....	25
2.2.1 Палеозойские отложения.....	25
2.2.2 Мезозойские отложения	26
2.2.2.1 Юрская система	26
2.2.2.2 Меловая система.....	29
2.2.3 Палеогеновая система.....	31
2.2.4 Четвертичная система.....	32
2.3 Тектоника.....	32
2.4 Нефтегазоносность	35
3 Анализ данных.....	37
4 Описание классификатора.....	42
5 Анализ результатов исследования.....	48
5.1 Анализ результатов решения задачи классификации	48
5.2 Анализ результатов выделения реперных горизонтов	53
6 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение.....	57
6.1 Анализ преимуществ реализации проекта.....	57
6.2 Планирование управления научно-техническим проектом	58
6.2.1 План проекта.....	58
6.2.2 Определение трудоёмкости выполнения работ	61
6.2.3 Бюджет научного исследования	64
6.2.3.1 Расчет заработной платы	64
6.2.3.2 Размер обязательных отчислений в фонды	65

6.2.3.3 Расчет подоходного налога	66
6.2.3.4 Расчет затрат на амортизацию специального оборудования	66
6.2.3.5 Расчет материальных затрат	67
6.2.3.6 Формирование бюджета затрат проекта	67
6.2.4 Реестр рисков проекта	68
6.3 Определение эффективности исследования	70
7 Социальная ответственность	73
7.1 Производственная безопасность	74
7.1.1 Вредные факторы	75
7.1.1.1 Освещение	75
7.1.1.2 Микроклимат	78
7.1.1.3 Ионизирующие излучения.....	79
7.1.1.4 Психосоматические аспекты работы.....	80
7.1.2 Опасные факторы	82
7.1.2.1 Поражение электрическим током	82
7.1.2.2 Статическое электричество	83
7.1.2.3 Короткое замыкание.....	84
7.2 Экологическая безопасность	85
7.3 Безопасность в чрезвычайных ситуациях	85
7.4 Правовые и организационные вопросы обеспечения безопасности	87
Заключение	90
Список публикаций студента.....	91
Список используемых источников.....	92
Приложение А	96

Введение

В современном мире многие, казалось бы, рутинные или напротив глобальные и нетривиальные проблемы решаются с помощью применения методов машинного обучения. От задач поиска новых созвездий или увеличения точности прогнозов валютных рынков до автоматизации работы продавца, везде теперь необходимо использование анализа и обработки данных. С увеличением количества информации человек становится неспособным обработать и проанализировать весь массив данных, а игнорируя или просто не замечая ряд признаков, приходит к неверным выводам и суждениям. Именно поэтому исследователям и специалистам необходимо использование методов анализа данных, в частности машинного обучения.

Машинное обучение позволяет произвести автоматизацию различных процессов, не просто увеличивая скорость обработки данных, но и повышая точность. Тем самым данные методы позволяют не только получить более точный результат за меньшее время, но и обнаружить в данных новые признаки.

Применение методов машинного обучения сейчас активно внедряется в нефтегазовом деле в различных аспектах. Успешно зарекомендовал себя такой метод машинного обучения как искусственные нейронные сети. Их применение помогает решать текущие технические проблемы, давать различные прогнозы по показателям работы месторождения, а также помогать решать задачи идентификации, возникающие на ранних стадиях жизни месторождения. Вышесказанное определяет актуальность настоящей магистерской диссертации.

Целью данной работы является применение искусственной нейронной сети для типизации геолого-геофизического разреза. Предметом анализа являлся набор каротажных данных с месторождений, территориально приуроченных к Нюрольской впадине.

Данная работа призвана уменьшить время, затрачиваемое специалистами геологами и геофизиками на анализ каротажных данных, и повысить точность производимой интерпретации.

Поставленная цель магистерской диссертации является нетривиальной и многомасштабной. Для достижения поставленной цели необходимо решить следующие задачи:

1. Обучение нейронной сети для точного выявления стратиграфии. Данная задача подразумевает вычленение крупных стратиграфических единиц на основе анализа поведения каротажных кривых с применением нейронных сетей.

2. Выявление реперных горизонтов внутри целевых стратиграфических единиц для дальнейшего более детального расчленения и корреляции. Поставленная задача может быть решена также с использованием алгоритмов машинного обучения.

3. Корреляция внутри обособленных реперными горизонтами интервалов. Данная задача имеет наименьший масштаб, как и поставленные ранее задачи может быть решена при использовании методов машинного обучения.

В рамках данной диссертации также был произведен анализ экономической эффективности внедрения проекта и рассмотрены основные вопросы, связанные с социальной ответственностью, возникающие при реализации данного исследования.

Защищаемые положения:

1. Применение нейронных сетей обоснованно для решения задачи типизации геолого-геофизических разрезов.
2. Классификатор на основе скрытых Марковских моделей позволяет эффективно решать задачу автоматического выделения реперных горизонтов.

1 Анализ предметной области и актуальность выполняемой работы

Аппарат нейронных сетей отлично зарекомендовал себя как инструмент для решения задач в нефтегазовой отрасли. Основным преимуществом применения нейронных сетей является то, что искусственные нейронные сети способны успешно справляться с задачами, осложнёнными большим количеством разнородной входящей информации.

Задача автоматизации процессов выделения в разрезе скважин интервалов горных пород активно решается рядом исследователей. В настоящий момент можно выделить несколько направлений использующих методы машинного обучения для решения данной задачи, а именно: выделение фаций; выделение крупных сегментов горных пород, соответствующих в стратиграфическом плане свитам; определение литологического состава отложений.

Метод выделения фаций с использованием машинного обучения рассмотрен в трудах как отечественных, так и иностранных исследователей. Так в работе [1] немецкие и французские специалисты: Tschannen, V. Delescluse, M., Rodriguez, M., и Keuper, J. решают задачу классификации фаций. Для решения данной задачи в работе применялся такая разновидность машинного обучения, как сверточные нейронные сети. Авторы использовали данные полученные от университета города Канзас. В их распоряжении находились данные по 11 скважинам: каротажные кривые и керновые данные. Материалы геофизических исследований состояли из каротажей: ГК, БК, НГК, ГГК-П.

Авторы произвели расчет нейтронной и плотностной пористости по данным НГК и ГГК-П и для обучения сети использовали их разницу, а также среднее значение [1]. Авторами была поставлена задача классификации девяти фаций таких как: песчаники континентального генезиса, грубозернистый и мелкозернистый алевролиты, образованные в континентальных условиях, аргиллит+алевролит морского генезиса. Также авторами выделены в

соответствии с классификацией Данхема (1962), следующие фации: мадстоун, вакстоун, пакстоун, отличие которых в пропорциях илистого матрикса и зернистого материала. Авторами также выделены чистые известняки и доломиты.

Нейронную сеть авторы исследования обучили на 9 скважинах и затем проверили на данных 2-х имеющихся скважин. В целом представленная в работе [1] сеть имела сходимость результатов предсказания фаций на уровне 70%. Анализируя полученный результат, авторы признают, что расхождение эталонного расчленения и полученного с применением нейросети вызвано разностью в масштабах методов. Специалисты геологи выделяли более тонкие пропластки, нейросеть анализировала лишь данные каротажных исследований, которые были нечувствительны к маломощным интервалам.

Аналогичную задачу выполняли исследователи политехнического университета Милана (Италия): Bestagini,P., Lipari,V.,Tubaro,S. и Milano,P. [2]. Использовался абсолютно такой же набор входных данных, как и в исследовании [1], те же каротаж и тот же набор из девяти фаций. Однако в данной исследовательской работе использовался такой алгоритм машинного обучения как «случайный лес». Результаты проделанных изысканий сравнивались с полученными результатами французско-немецкой команды исследователей. В целом данное исследование показало, что лишь часть фаций была успешно опознана с использованием алгоритма «случайный лес». Так, например, предсказанный грубозернистый алевролит континентального генезиса имел сходимость с реальными данными 77%, однако фация мадстоуна совпала лишь на 12,8%.

В работе [3] чешские специалисты: Malvić, T., Velić, J., Horváth, J., и Svetković, M. анализировали данные с трех месторождений, расположенных на территории Чешской республики. В отличие от [1] для реализации той же цели не использовались сверточные нейронные сети, а использовались нейронные сети прямого распространения. В статье [3] рассмотрен ряд алгоритмов работы нейронных сетей примененных для реализации задачи, а именно:

вероятностные нейронные сети, сеть радиально-базисных функций и метод эластичного распространения ошибки.

Вероятностные нейронные сети, используемые для задач классификации, являются разновидностью байесовских сетей. Такие сети применяют оценку плотности вероятности по уже имеющимся данным, при этом используя ядерные оценки. По сути, делается допущение, что уже существующая точка наблюдения в пространстве создает некое значение плотности вероятности для данной области. Области пространства, где точки наблюдения расположены особенно близко, следовательно, имеют высокое значение плотности вероятности, представляют собой кластеры. Оценка доверия к уровню плотности вероятности осуществляется на основе расстояний от ядерных центров [4].

Радиально-базисные нейронные сети используют функцию активации радиального типа. Особенностью сетей такого типа является отсутствие необходимости в большом количестве скрытых слоев. В целом достаточно одного слоя. Выходной слой сети состоит из одного или нескольких линейных нейронов.

Метод эластичного распространения ошибки является аналогом метода обратного распространения ошибки, однако анализируется не сама частная производная функции ошибки, а лишь ее знак. Это обеспечивает сокращение времени необходимого для обучения сети.

В работе [3] использовались следующие каротажные кривые: ГК, ГЗ, кривая глинистости, общая пористость, к сожалению, авторами не указана природа каротажей, по которым были определены значения пористости и глинистости. Авторы указывают, что также исследовали два варианта обучающей выборки. Первая набор из трех каротажных кривых, вторая выборка из 7-ми. Точность предсказания составила 78,3% для меньшей выборки и 82,1% для большей.

Большой интерес для решения задачи классификации представляет работа Y.Zee Ma. [5]. В данном исследовании ставится цель найти метод

машинного обучения, который позволил бы достоверно отделять фации песков разлива, фации дельтовой равнины и фацию русловой отложений. Автор имел в распоряжении данные: ГК, ПС, ГГК-П и НГК, особенностью проведенного исследования является применение нейронного алгоритма анализа главных компонентов, а именно использование каскадной конфигурации сетей. Использование каскада из нескольких более простых сетей позволяет уменьшить время обучения, по сравнению с одной более сложной по конфигурации сетью.

Задача типизации геолого-геофизического разреза также имеет примеры реализации. Так Сет Уиллис Бассет из геологической службы штата Флорида, США использует методы машинного обучения для определения формаций [6]. Также в работе выделяются реперные горизонты, строятся карты в автоматическом порядке. В данной работе автор рассматривал такой вид нейронных сетей, как машины опорных векторов и сравнивал точность его результатов с такими алгоритмами как дерево решений и случайный лес. В данном исследовании использовались данные со 144 скважин, в которых было известно значения ГК и истинной глубины. Несмотря на использование такого ограниченного объема данных, результаты имели точность достигающую 95%.

Схожую задачу решают Gupta I., Rai C.S., Sondergeld C.H., и Devegowda D. по данным с нескольких месторождений в Вудфордской нефтегазоносной провинции [7]. Анализировалось 12 скважин, входной набор данных представляли собой: ГК и ИК. Также как и в работе [6] выделялись как формации, так и реперные горизонты. В данной работе применялись: машины опорных векторов, метод k -средних, а также метод самоорганизующихся карт Кохонена.

Особенного внимания при анализе методов решения задачи типизации геолого-геофизического разреза заслуживает работа Шайбакова Р.А. [8]. В данном исследовании рассматривается вопрос реализации типизации разреза, с применением метода самоорганизующихся карт Кохонена. Анализ подверглись данные с 12 скважин Правдинского месторождения,

расположенного в Ханты-Мансийском автономном округе. Исходными данными являлись: ГК, ПС, ИК, а также значения абсолютных отметок. Результатом работы является получение реперных горизонтов, которые на 64% совпадали с маркерами определенными профессиональным интерпретатором. Автором подчеркивается, что для корректной реализации типизации геолого-геофизического разреза обязательным условие является определение обстановки осадконакопления, для понимания возможности корреляции тел.

Задача определения формаций с помощью машинного обучения рассматривается в работе Bougher, B.C. [9]. Автор приведенной научной статьи ставит задачу классификации 8ми формаций на основе анализа каротажных данных. В рамках данного исследования использовался метод k -ближайших соседей. Ввиду малого количества входных параметров, точность определения формаций составила в среднем лишь 36%, для отдельных формаций точность достигала 67%.

Комплексная работа по интерпретации геофизических данных, проведена Гафуровым Д.О., Гафуровым О.М. и Конторовичем В.А. в статье [10] на основании геофизических данных Талаканского нефтегазоконденсатного месторождения. Важной особенностью является анализ представительности выбранной входной информации. Так авторами анализировались результаты бти выборок каротажных данных, наиболее объемная содержала 7 каротажей, а наименьшая лишь 3. Так наименьшая выборка содержала: ГК, НГК, АК и показала результат в 67.5% правильных определений. Замена акустического каротажа в обучающей выборке на ДС увеличила точность получаемых результатов до 78.4%. Использование совокупности из 4х каротажей: ГК, НГК, АК, ДС повысили точность до 94.9%. Последовательное добавление данных геофизических исследований методами: БК; МБК; ГГК-П привело к увеличению точности до 99.6%. Однако стоит отметить, что составить выборку из скважин обладающими полным набором каротажных исследований является практически невыполнимой задачей для представленного для анализа изучения. Поэтому дальнейшие исследования,

представленные в данной магистерской диссертации, опираются на опыт приведенный в работе [10] касаясь части составления обучающей выборки.

Актуальность темы магистерской диссертации подчеркивается пока еще не полной изученностью данной сферы применения методов машинного обучения. Сами методы машинного обучения разработаны уже давно, но их прикладное применение в нефтегазовой области для целей анализа данных геолого-геофизических исследований находится еще на ранней стадии. Применение такого метода машинного обучения, как нейронные сети для задачи типизации разреза имеет существенное преимущество перед классической интерпретацией специалистом-геологом в том, что длительность процесса интерпретации сокращается значительно. Таким образом при наличии большого количества скважин, это имеет значительный и экономический эффект сокращая трудозатраты специалистов и ускоряя процесс дальнейшей корреляции, что является актуальной целью для нефтегазовых предприятий.

Также использование нейронных сетей для целей типизации геолого-геофизического разреза на месторождениях Ньюрольской впадины ранее не применялось. Следовательно, данная работа в рамках данной тектонической структуры является уникальной, а полученные результаты ее выполнения могут быть использованы в качестве сравнительной базы для дальнейших работ в данной сфере.

2 Анализ геологической изученности района исследований

2.1 Географическое расположение и изученность территории

Исходными данными данного проекта являлись геофизические исследования скважин, предоставленные с трех месторождений (В, С, D), структурно приуроченных к Нюрольской впадине (Рисунок 2.1). Месторождения административно приурочены к Парабельскому району томской области. Парабельский район расположен в центральной части Томской области и простирается с юго-запада на северо-восток. На севере он граничит с Курганской областью, на востоке – с Верхнекетским, на юге и юго-западе – с Колпашевским и Бакcharским районами, на западе с Новосибирской областью.

В дальнейшем рассмотрим основные вопросы, связанные с геологическим строением и физическими характеристиками на примере месторождения С.

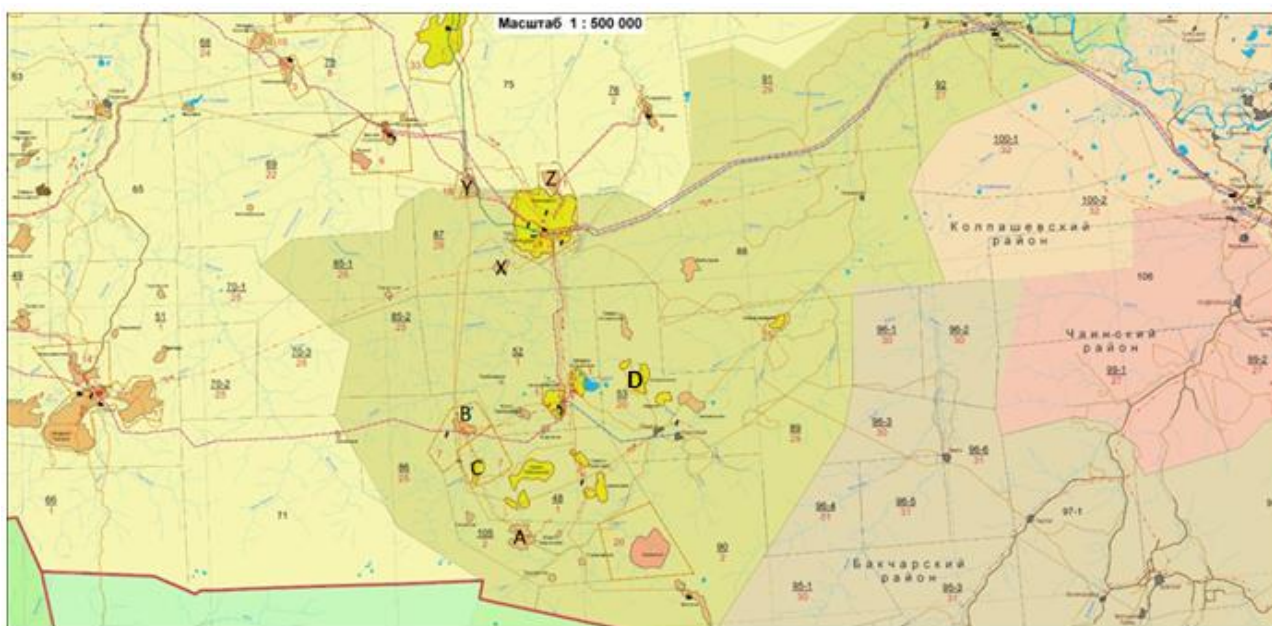


Рисунок 2.1 – Обзорная карта исследуемой области

Месторождение В открыто в 1974 году, приурочено к эрозионно-тектоническому выступу Нюрольской впадины. По типу флюида является нефтегазоконденсатным. Промышленная залежь приурочена к зоне контакта дезинтегрированных пород палеозойского возраста и юрских отложений.

Залежь относится к структурно-литологическому типу и является тектонически экранированной.

Месторождение С открыто в 1985 году, приурочено к юго-западному склону Нюрольской впадины в зоне сочленения ее с Лавровским валом. По типу флюида является нефтегазоконденсатным. Продуктивными являются позднепалеозойские отложения, а месторождение приурочено к Пудинскому нефтедобывающему району.

Месторождение D открыто в 1967 году, приурочено к структурному мысу, осложняющему юго-западную периклиналь Юбилейного поднятия в зоне его сочленения с юго-западным погружением Пудинского мегавала. По типу флюида является нефтяным.

Месторождения окружены таежным массивом и заболоченной местностью, потому территория, на которой расположены месторождения, является труднопроходимой. Через территорию лицензионных участков месторождений протекают реки Чижалка, Чузик, Армич, Язевка и Ершовка обладающие большим количеством притоков. Северная часть лицензионных участков представляет собой непроходимое болото. Основная часть площади месторождений покрыта смешанным лесным массивом.

В климатическом отношении район имеет резко континентальный климат и подвержен влиянию арктических воздушных масс. Благодаря чему состояние погоды неустойчиво и происходят частые колебания температуры и количества выпадающих осадков. Средняя измеренная годовая температура составляет 1,5°C. Минимальная среднемесячная температура воздуха приходится на январь и составляет -20,6°C, а наименьшая зарегистрированная температура в этот месяц составляет -55°C. Длительность безморозного периода составляет 89 дней в году. Наиболее теплыми месяцами являются июнь и июль, температура воздуха достигает отметок в +35°C.

С 1963 года в рассматриваемом регионе проводились крупно- и среднемасштабные высокоточные гравиметрические исследования. В

комплексе с другими методами также выполнялись электроразведочные работы [11].

С 1981 по 1982 года на территории месторождений проводились сейсмические изыскания, была определена структура и выявлены признаки наличия углеводородов. На этом основании началось разведочное бурение площади. Так на месторождении С проект поискового бурения предусматривал бурение 7 поисково-разведочных скважин. Объектами интереса стали отложения палеозоя, а также песчаные тела васюганской и тюменской свит. В результате реализации проекта были пробурены скважины (40, 41, 43, 44, 45, 46, 51), которые вскрыли продуктивные отложения палеозойского фундамента.

В результате бурения было установлено, что породы, слагающие мезозойский чехол залегают практически горизонтально, а карбонатные породы, слагающие палеозойский фундамент залегают наклонно, с углом, достигающим 50 градусов. Отбор кернового материала из продуктивной части разреза выполнен в малых количествах, в большей части отобраны полные карбонатные породы.

2.2 Геологическое строение территории исследования

2.2.1 Палеозойские отложения

По рассматриваемой зоне известно, что органогенные постройки верхнего девона развиты в виде карбонатных банок, биостромов и биогермов по перифериям поднятий вдоль уступов по бортам Нюрольской впадины, осложненной разломами. Здесь селились бентосные формы организмов (фораминиферы, строматопориты, кораллы, водоросли). По латерали в пределах депрессионных зон они замещались карбонатно-кремнистыми битуминозными осадками доманикового типа, включающими фораминиферы, радиолярии, мшанки, остракоды, тентакулиты, конодонты. Верхнедевонский этап характеризовался формированием биогермных известняков, строматопорово-водорослевых, сферово-узорчатых, органогенно-обломочных, каверново-порового типа [12].

Переотложения верхнефаменских карбонатов обусловлены процессами размыва в турнейском веке. В научной литературе часто встречается рифовая модель строения карбонатных массивов верхнего девона Урманско-Арчинской зоны. В литологическом плане отложения представлены органогенными известняками, амфипоровыми, строматопоровыми, светловато-серыми, бурыми, кавернозными и трещиноватыми. Трещины известняков заполнены кальцитом от белого до розоватого. В кровле отмечено развитие трещин, зон каолинизации, кремнения, сидеритизации. Под влиянием вторичных процессов осадконакопления происходило образование пустотного пространства.

2.2.2 Мезозойские отложения

2.2.2.1 Юрская система

На карбонатных породах фундамента залегают отложения юрской системы с угловым несогласием, сформированное при перерыве в осадконакоплении. Юрские отложения представлены тюменской, васюганской, георгиевской и баженовской свитами.

Отложения формирующие тюменская свиту в районе исследования, в основном, являются породами озёрно-аллювиального и озёрно-болотного генезиса, т.е. континентальными осадками, в отличие от васюганкой, георгиевской и баженовской свит, имеющих преимущественно прибрежно-, мелководно- и глубоководно-морских условиях. Отложения тюменской свиты характеризуются, большой фациальной и литологической изменчивостью по площади и по разрезу. Прослаивание песчано-алевролитовых разностей, аргиллитов и углей. Песчаники, слагающие тюменскую свиту имеют от светло-серого до серого окрас, имеют полимиктовый состав, часто являются глинистыми и известковистыми. Аргиллиты от серых до черных, крепкие, углистые, отмечена сидеритизация. Мощность отложений тюменской свиты составляет от 250 до 400 м. Сейсмический отражающий горизонт I-а приурочен к подошве тюменской свиты, а горизонт I-б к ее кровле.

Верхняя граница васюганской свиты по данным палеонтологических исследований соотносится с оксфордским ярусом, подошва – батским.

В литологическом плане васюганская свита подразделяется на нижнюю и верхнюю васюганскую подсвиту.

Нижневасюганская подсвита в основном состоит из аргиллитоглинистых разностей, имеющих цвет от тёмно-серых до чёрных, буроватых, иногда с зелеными оттенками, с редкими малочисленными прослойками алевролитов. В основании подсвиты принято выделять пахомовскую пачку, которая сложена в основном слабо отсортированными песчано-алевролитами с остатками морских организмов и следами их жизнедеятельности. В целом нижневасюганская подсвита имеет мощность в 30-40 метров, однако может достигать 60 метров в местах наибольшего погружения палеорельефа. Согласно исследованиям, А. С. Турбиной возраст отложений определен как позднекелловей-раннеоксфордский по остаткам пелецпод *Meleagrinea*. Спорово-пыльцевой анализ, проведенный О.Н. Костеша (ТГУ) указывает на средне-позднекелловейский возраст отложений.

Верхневасюганская подсвита состоит из песчаников и алевролитов светло-серых, буроватых, с прослоями аргиллитов. Полный геологический разрез верхней части Васюганской свиты содержит 4-5 песчаных пластов.

Так по наличию регрессивного и трансгрессивного циклов во время формирования верхневасюганской подсвиты делает возможным подразделить ее на две пачки - подугольную и надугольную.

Минимальному уровню моря или максимуму регрессии соответствует регионально-выдержанный угольный пласт $У_1$, сформированного в континентальных условиях осадконакопления.

Подугольная пачка состоит из регрессивных песчаных пластов $Ю_1^3$ и $Ю_1^4$, часто эти пласты объединяются в один $Ю_1^{3-4}$. Отложения подугольной пачки содержат остатки раковин пелецпод, следы жизнедеятельности морских донных организмов.

Надугольная пачка расположена между подошвой георгиевской свиты и угольным пластом У₁. Характер распространения песчаных тел надугольной пачки обусловлен обстановкой осадконакопления. Мощность надугольной пачки увеличивается по направлению к палеовпадинам. На месторождении С в нижней части отложений надугольной пачки обнаружено присутствие флоры: *Coniopteris latilobus Bistr.*, *C. simplex (L. et. H.) Harris.*, *C. maakiana Prindata C. Ex gr. Burejensis*, *Pityophyllum sp.* По остаткам моллюсков *Dentallina sp.* возраст отложений определен, как раннеоксфордский.

Георгиевская свита залегает между васюганской и баженовской свитами. В литологическом плане отложения свиты представляют собой глинисто-аргиллитовую разность, редко слабобитуминозную. Также присутствует алевролитовый материал и зерна глауконита. Мощность отложений георгиевской свиты изменчива по латерали. Отложения могут отсутствовать, а могут достигать 20 метров. В отложениях георгиевской свиты отмечаются многочисленные остатки макрофауны (аммониты, белемниты, двустворки). Комплекс микрофауны представленный в породах свиты позволяет достоверно определить кимериджский возраст отложений. В западной части Томской области георгиевская свита представлена фрагментарно, основной район ее накопления – Нюрольская впадина. Относительно широкая полоса, в пределах которой получила развитие георгиевская свита протягивается вдоль южного обрамления Западно-Сибирского бассейна в зоне, где отложения баженовской свиты становятся низкоомными [11].

Баженовская свита трансгрессивно, с несогласием залегает на отложениях верхневасюганской подсвиты или согласно перекрывает аргиллиты георгиевской свиты. Толща хорошо выдержана по латерали. Нижняя граница свиты проведена в нижней части волжского яруса, верхняя - в низах нижнего берриаса.

Породы содержат обильную морскую фауну: ихтиодетрит, радиолярии, фораминиферы, ростры белемнитов, раковины пелеципод. В палинокомплексе отложений доминирует микрофитопланктон *Acritarha*.

Баженовская свита в литологическом плане представлена глинисто-кремнисто-карбонатными породами, черными, битуминозными. Часто породы имеют коричнево-черный оттенок, аргиллиты массивные, плитчатые. В отложениях отмечены прослои радиоляритов, глинистых известняков, широко распространены остатки морских фоссилей. Содержание органического вещества может достигать более 20% в высокоуглеродистых аргиллитах. Баженовская свита является нефтематеринской свитой для осадочного чехла Западно-Сибирского нефтегазоносного района. Мощность битуминозных баженовских аргиллитов составляет в среднем от 10 до 30 метров.

Отличительной особенностью баженовской свиты является ее резкое отличие в физических свойствах как от выше- и нижележащих отложений, так и от мезозойско-кайнозойского разреза. Битуминозные глинисто-кремнисто-карбонатные породы с легкостью выделяются на каротажах. Породы обладают высокими значениями кажущего сопротивления, достигающих 500 Ом·м, также обладают высоким уровнем естественной гамма-активности. Породам баженовской свиты также характерно высокие значения нейтронного гамма-каротажа, а за счет углеводородов породы обладают высоким электрическим сопротивлением. Однозначность определения баженовской свиты по различным методам сделало породы, слагающие данную свиту наилучшим геологическим репером на исследуемой территории. Также отложения баженовской свиты являются реперными горизонтами для сейсмических изысканий, так подошва баженовской свиты относится к отражающему горизонту Па, а ее кровля к Пб.

2.2.2.2 Меловая система

Куломзинская свита залегает согласно на отложениях баженовской свиты в литологическом плане представлена аргиллитами с переслаиванием

алевритов и аргиллитов. В подошве свиты выделяется – ачимовская пачка, представленная известковистыми песчаниками темно-серого цвета. Ачимовская пачка, как в целом куломзинская свита является региональным флюидоупором, препятствующим миграции углеводородов, генерируемых баженовской свитой вверх по разрезу. Мощность отложений свиты в среднем составляет 250-300 м. возраст отложений определен по остаткам морской фауны, как берриас – частично ранний валанжин.

Тарская свита в литологическом плане представлена мелкозернистыми песчаниками от светло-серых до серых. Песчаники полимиктовые, в состав матрицы породы входят кварц и полевые шпаты, цемент гидрослюдисто-хлоритовый. Тарская свита сформировалась в прибрежно-морских условиях. Мощность тарской свиты в среднем 90-110 м.

Киялинская свита сложена континентальными отложениями, которые согласно залегают на породах тарской свиты. В литологическом плане породы киялинской свиты представляют переслаивание глин, алевритов и песчаников, с преобладанием в составе отложений глинистого материала. Киялинская свита сформировалась в переходных условиях, наиболее вероятным представляется наличие лагунной обстановки седиментации. В отложениях свиты замечена карбонатизация. Мощность свиты на месторождении С составляет в среднем 640-680 метров.

Алымской свита залегает на отложениях Киялинской свиты. Породы алымской свиты имеют переходный от прибрежно-морского до морского генезис. Отложения свиты подразделяются на нижнюю и верхнюю подсвиты. Нижняя подсвита сформирована песчаными отложениями, имеющими серый и светло-серый оттенок. Песчаник полимиктовый, матрица породы содержит полевые шпаты и кварц. Верхняя часть алымской свиты представлена породами, имеющими морской генезис. Верхнюю часть свиты составляют темно-серые и черные каолинистые глины. Мощность свиты на месторождении С достигает 50 м.

Покурская свита залегает на отложениях Алымской свиты. Свита состоит из пород континентального и частично прибрежно-морского генезиса. Отложения в литологическом плане представлены серыми и зеленовато-серыми песчаниками. В керновом материале отмечены прослои мергелей, глинистых известняков, а также включения растительного детрита. Мощность отложений свиты в пределах площади месторождения С практически не изменяется и составляет порядка 880 метров.

Кузнецовская свита в литологическом плане представлена серыми и темно-серыми плотными аргиллитами. Отложения кузнецовской свиты имеют морской генезис. Выделены следующие морские организмы *Baculites romanovskii* Arkh., *Inoceramus ef. labiatus* Schloth. и фораминиферы с *Gaudryina filiformis* Berth. Для пород, слагающих кузнецовскую свиту, характерно наличие аутигенного пирита, глауконита, а также фосфатных минералов. Мощность отложений свиты достигает 50м.

Отложения ипатовской свиты представлены переслаиванием песчаников и аргиллитов, имеющих прибрежно-морское происхождение осадка. Мощность отложений на месторождении С достигает 200 метров.

Славгородская свита представляет собой зеленовато-серые аргиллиты морского генезиса, содержит остатки фораминифер и радиолярий. Мощность отложений славгородской свиты на месторождении С до 70 метров.

Ганьская свита залегает на отложениях Славгородской свиты согласно. В литологическом плане отложения ганьской свиты представлены переслаиванием аргиллитов серых, с тонкими прослоями серых и темно-серых алевролитов, местами отмечены мергели. В отложениях ганькинской свиты отмечены остатки фораминифер и остракод. Мощность отложений ганьской свиты вскрытой на месторождении С составляет до 100 метров.

2.2.3 Палеогеновая система

Палеогеновая система включает в себя отложения талицкой (палеоцен), люлинворской (эоцен), чеганской (верхний эоцен - нижний олигоцен) свит,

имеющих морской генезис осадков их слагающих, также включает в себя отложения некрасовской серии (средний - верхний олигоцен), имеющие континентальный генезис, и которые согласно залегают на породах слагающих меловую систему. Совокупная мощность отложений палеогена составляет на месторождении С до 350 метров.

Талицкая свита в литологическом плане представлена темно-серыми и черными плотными аргиллитами, отмечены мелкие пропластки алевролита, а также включения пирита. В отложениях свиты встречаются остатки фораминифер и остракод.

Люлинворская свита представлена зеленовато-серыми аргиллитами с редкими прослойками серых слюдистых алевролитов и слабосцементированных песчаников. В отложениях свиты отмечено присутствие радиолярий.

Чеганская свита на месторождении С представлена в литологическом плане аргиллитами голубовато-зеленого цвета, местами зеленовато-серыми, плотными. Также отмечены линзовидные прослои песчаника серого полимиктового, в составе матрицы которого встречаются полевые шпаты и кварц.

2.2.4 Четвертичная система

Отложения четвертичной системы на месторождении С представлены серыми песками мелко-среднезернистыми, суглинками и буровато-серыми глинами, с пропластками лигнита и почвенно-растительным слоем. Общая мощность отложений четвертичной системы составляет в среднем 30м.

2.3 Тектоника

В строении рассматриваемого района участвуют фундамент герцинского возраста и мезозой-кайнозойский осадочный чехол (мощностью 3 км) с двумя этажами нефтегазоносности. Нижний (доюрский) этаж связан с залежами, обнаруженными в породах переходного комплекса и складчатого основания.

Верхний этаж нефтегазоносности (юрско-кайнозойский) образован породами осадочного чехла, накопившимися в рифтах, надрифтовых прогибах и депрессиях. Большинство нефтепроявлений установлены на контакте мезозойских (юрских или меловых) и доюрских пород различного состава (в коре выветривания) – продуктивные пласты М, М1 [13].

По отложениям нижнего структурного этажа фундамента, месторождение С относится к Межовскому срединному массиву (Рисунок 2.2), который представляет крупный выступ, сформированный байкальской складчатостью, который был переработан в герцинскую эпоху. Межовский массив с запада граничит с Колтогорско-Уренгойским грабен-рифтом, а с севера и северовостока с Нюрольским прогибом.

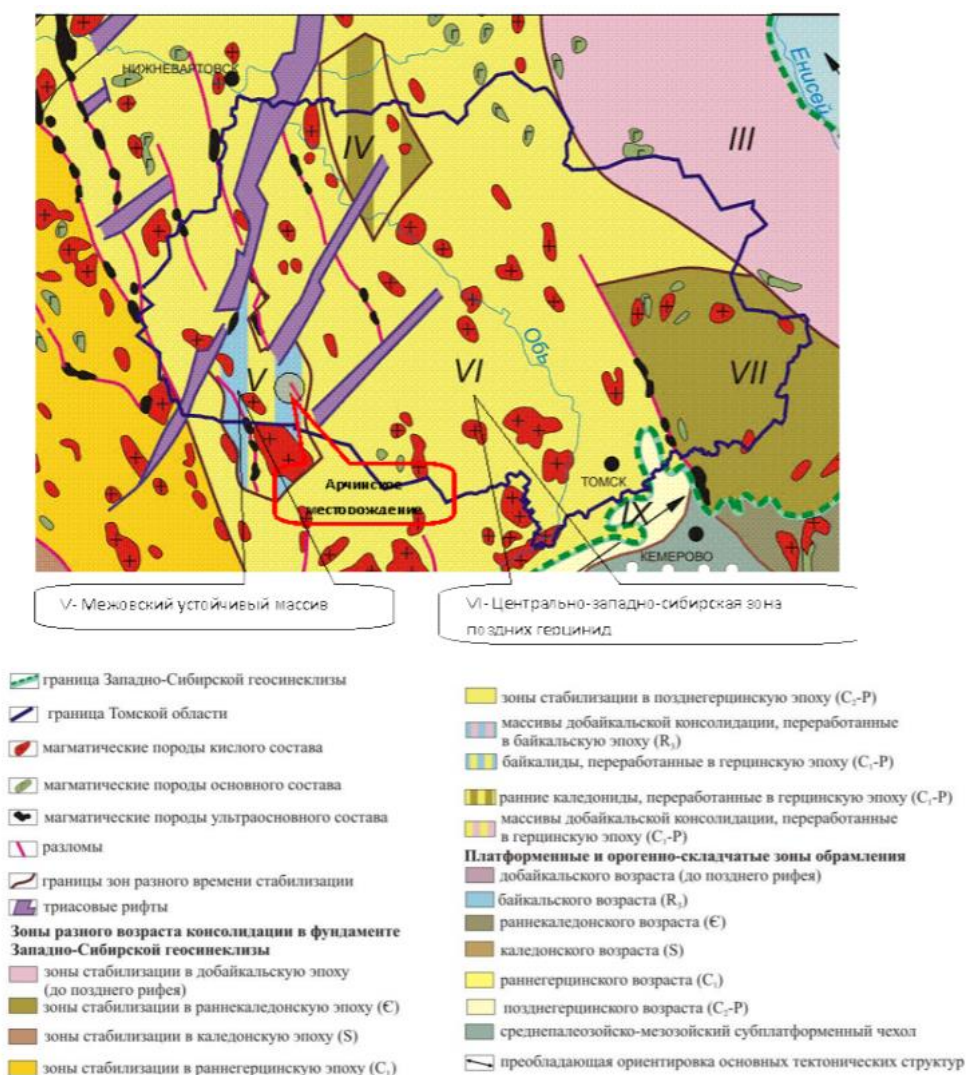


Рисунок 2.2 – Фрагмент тектонической карты фундамента Западно-Сибирской платформы (Сурков В.С., Жеро О.Г.), 1981г.

Наличие тектонических нарушений установлено на месторождении С благодаря данным сейсмических исследований, а также проведению гидродинамических исследований. В работе [13] приводятся новые данные о тектонических нарушениях на исследуемой площади месторождения С.

На площади месторождения выявлено три типа – разрывных нарушений: северо-восточного, субширотного и северо-западного простираний (Рисунок 2.3). По протяженности, устойчивости знака сдвига (правый или левый сдвиг) и количеству определений осей сжатия и растяжения. Согласно результатам проведенных сейсмических работ, в границах площади месторождения С наблюдается преимущественно сжатие северо-северо-западного направления, также отмечены единичные субширотные ориентировки. Подтверждено наличие проницаемых и непроницаемых разломов.

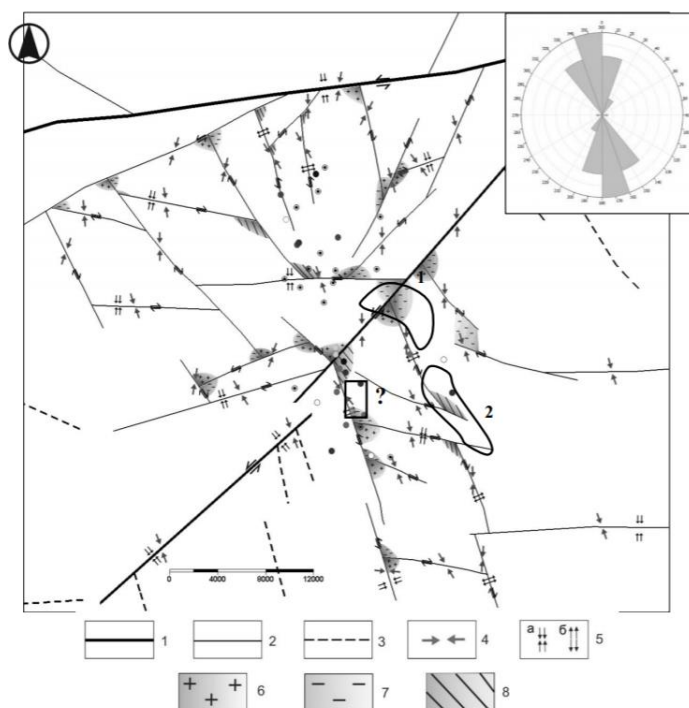


Рисунок 2.3 – Схема новейшей разломной тектоники и неотектонических напряжений на месторождении С [13]

Дебиты скважин отмечены оттенком серого (от светлого – скважины с низким дебитом, до темного – скважины с высоким дебитом). Условные обозначения: 1 – Разломы I ранга; 2 – Разломы II-ранга; 3 – Предполагаемые

разломы; 4 – Проекция ориентировки оси сжатия σ_3 ; 5 – Обстановка дополнительного сжатия (а) или растяжения (б), нормальная к плоскости разлома; 6 – Участки локального сжатия; 7 – Участки локального растяжения; 8 – Участки неполного растяжения

Анализ сейсмического материала проведенной в работе [14] позволил выделить на месторождении С ряд разломов приведенных на Рисунке 2.4.

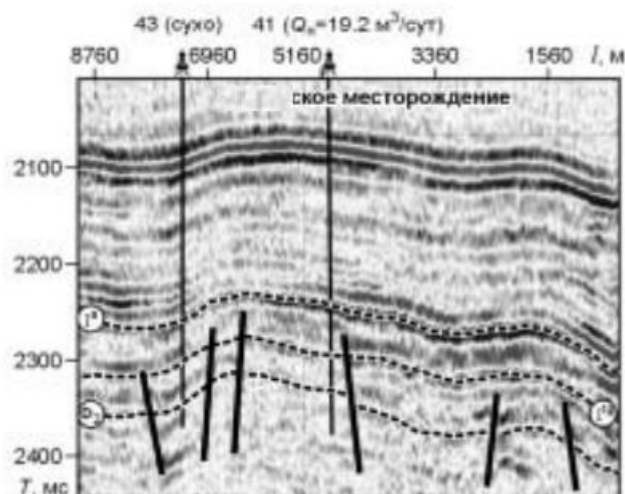


Рисунок 2.4 – Временной разрез по данным месторождения С

2.4 Нефтегазоносность

На месторождении С, нефтяная залежь массивного типа с газовой шапкой, приурочена к карбонатному выступу палеозойского фундамента.

В работе [14] резюмируется, что наиболее перспективными в плане нефтегазоносности коренных отложений палеозоя, являются эрозионно-тектонические выступы, которые в литологическом плане сложены органогенными известняками и не перекрыты корой выветривания.

Кровля залежи на месторождении С отмечена на глубине 3011м, ее подошва принята на глубине 3119м. Мощность залежи в сводовой структуре подтверждена данными скважины №41. На границе структуры мощность уменьшена до 13 метров, что подтверждено данными скважины №46. Газонефтяной контакт установлен на абсолютной отметке 2940 м, как наиболее высокая точка с полученным притоком нефти, в скважине №40. Принятое

положение ГНК согласуется с данным полученными в результате испытаний других скважин месторождения С. Водонефтяной контакт для залежи на месторождении С установлен на абсолютной отметке в -3002 м., согласно испытаниям скважины № 41.

Тип коллектора залежи месторождения С является порово-трещиноватым, местами кавернозным. Карбонатные породы-коллекторы обладают низкой механической прочностью, и редко могут быть извлечены на поверхность при процедуре отбора кернового материала. Поэтому, для выделения эффективных мощностей и определения подсчетных параметров использовались данные промыслово-геофизических исследований. Сложное геологическое и тектоническое строение палеозойских отложений фундамента обуславливает сложность гидродинамического моделирования поведения флюида в данном типе резервуара.

Строение залежи месторождения С показано на Рисунке 2.5, красными линиями цветом отмечены тектонические нарушения.

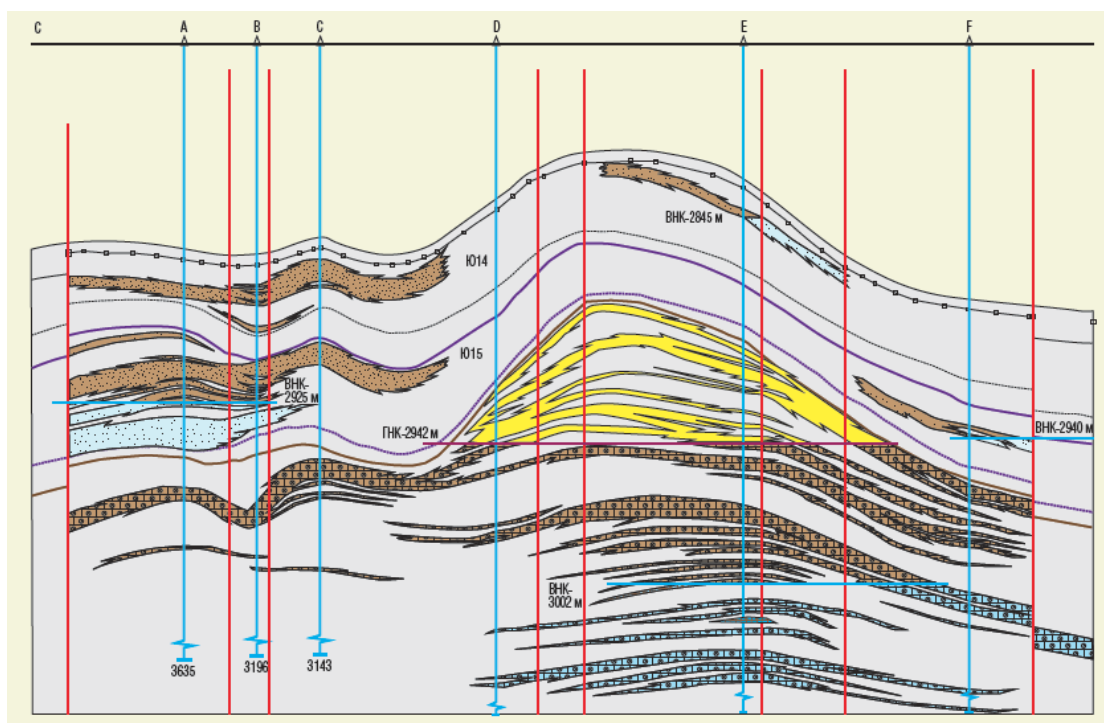


Рисунок 2.5 – Схема строения залежи месторождения С [15]

3 Анализ данных

Сырые входные каротажные данные каждой скважины хранятся в файлах формата .las. Данные одинаковых методов каротажных измерений в соседних скважинах могут иметь значительно отличающиеся средние значения ввиду различных условий проведения измерений, а также содержат в себе пропуски и могут быть прописаны в различных диапазонах. Перечислим основные проблемы, имеющиеся в исходных данных:

- различный набор каротажных данных;
- большое количество пропусков в каротажах;
- малое количество данных на малых глубинах, большое количество в зонах интереса, неравномерное распределение классов;
- различная длина участков однотипных каротажей одинаковых свит для разных скважин;
- различные условия проведения измерений (например, по-разному откалиброванными приборами), вследствие чего данные в одной и той же точке имеют разные числовые показатели;
- шумы и выбросы в данных.

На первом этапе подготовки данных по каждой скважине отдельно составляются каротажные выборки с отброшенными пропущенными или неверными (меньше нуля) значениями. Затем для решения проблемы различающихся условий измерений проводится нормировка каротажей путем вычисления стандартизированных оценок (z -оценок). Стандартизованная оценка является безразмерным статистическим показателем для сравнения значений разных размерностей или шкал измерений.

Основываясь на предположении о нормальности распределения данных, множество данных каротажа $X = \{x_1, x_2, \dots, x_n\}$ со средним значением \bar{X} и стандартным отклонением S_x можно преобразовать в другое множество с нулевым средним значением и единичным стандартным отклонением таким образом, что преобразованные значения z будут непосредственно выражаться в

отклонениях исходных значений от среднего, измеренных в единицах стандартного отклонения. Таким образом, каждая точка каротажа преобразуется по следующей формуле: $x_i = \frac{x_i - \bar{X}}{S_x}$.

Приведем полученные гистограммы стандартизованных оценок каротажей по данным четырех скважин месторождения С (Рисунок 3.1 – 3.3).

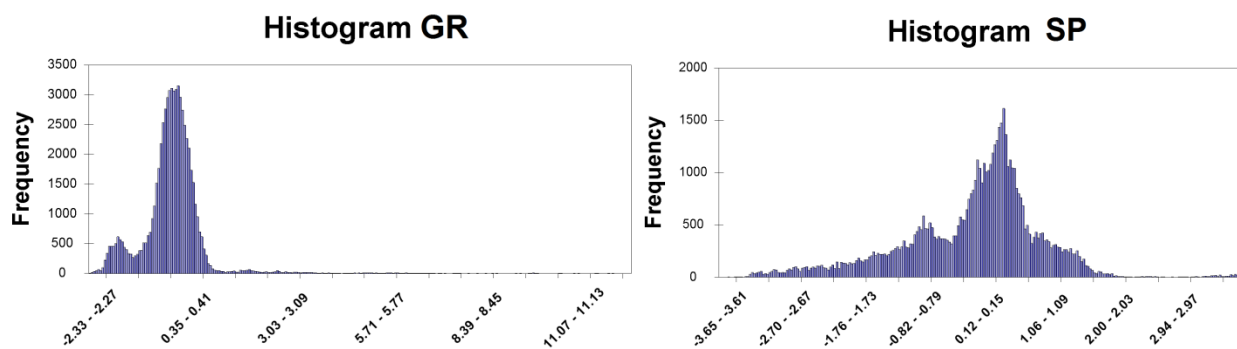


Рисунок 3.1 – Гистограммы стандартизованных данных каротажей гамма и собственной поляризации

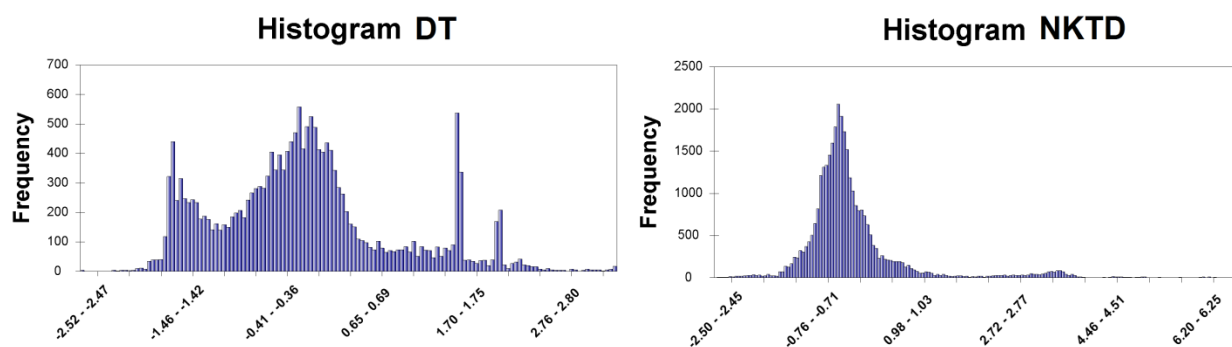


Рисунок 3.2 – Гистограммы стандартизованных данных акустического и нейтронного гамма каротажей

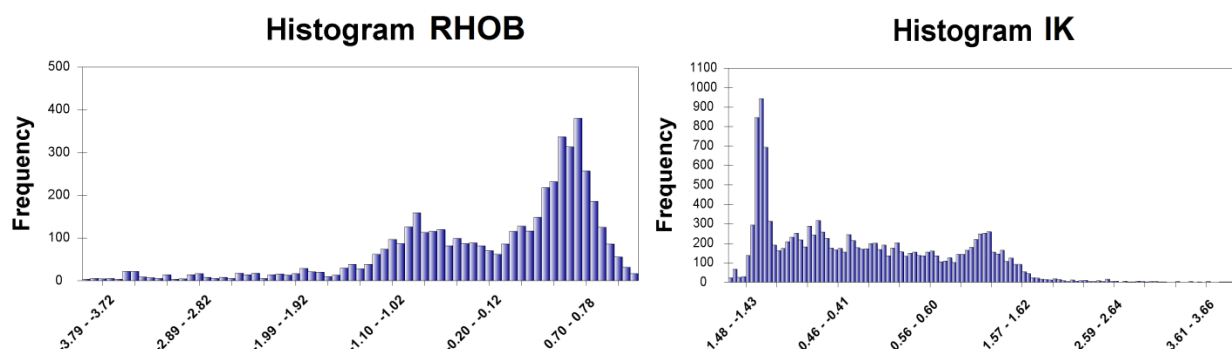


Рисунок 3.3 – Гистограммы стандартизованных данных плотностного и индукционного каротажей

Визуализация исходных данных на этапе разведывательного анализа помогает лучше понять исходные закономерности в данных: так, например, распределение индукционного и плотностного каротажей имеют логнормальный вид. Преобразование входных данных в соответствии с распределениями позволяет улучшить качество и уменьшить скорость обучения нейронных сетей.

На Рисунке 3.4 приведена гистограмма распределения количества точек входных данных по классам (свитам).

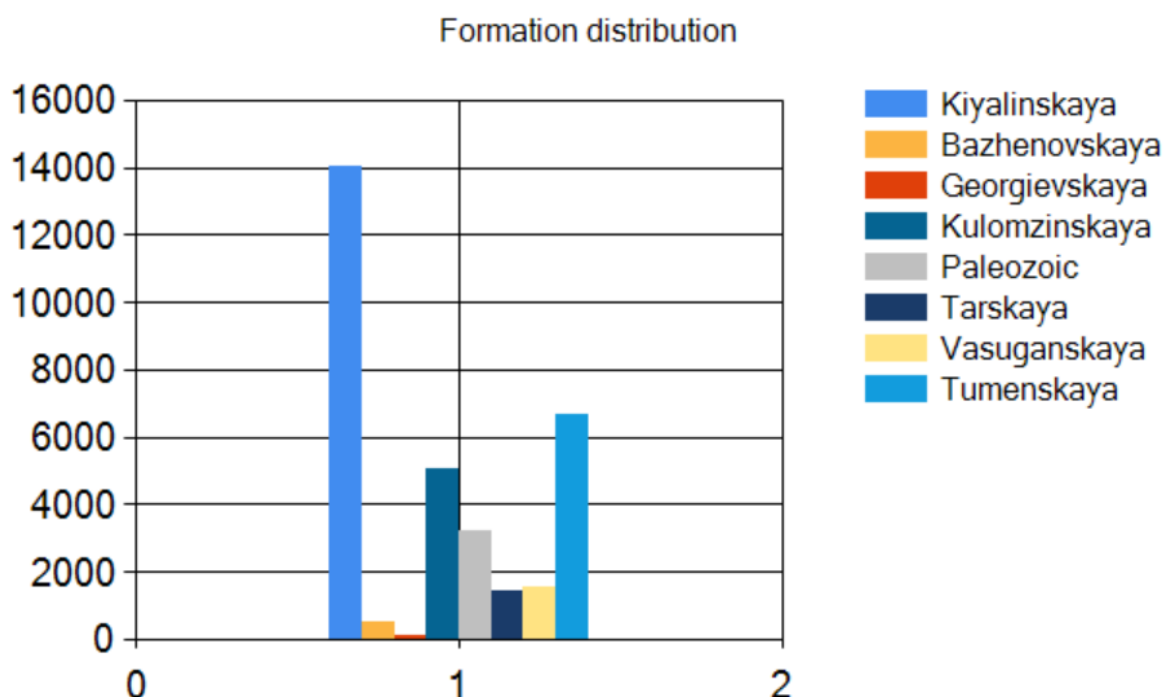


Рисунок 3.4 – Гистограмма распределения данных по свитам

По Рисунку 3.4 виден еще один недостаток исходных данных – различное количество обучающих векторов по классам. Для решения данной проблемы используется перекрестная энтропия в качестве критерия оптимизации, которая уменьшает меру неопределенности классификатора.

Измерения каротажей приводятся по точкам с фиксированным шагом глубины, однако он может различаться для разных скважин. При составлении обучающей выборки каротажные данные должны быть дискретизированы по абсолютной глубине. В качестве шага дискретизации используется значение в

0,1 м. Чем детальнее шаг дискретизации, тем детальнее будет выглядеть картина геологического профиля месторождения.

При решении задач классификации выделяют три основных типа задач:

- бинарная классификация (положительный и отрицательный признак).
- многоклассовая классификация (объект имеет признаки исключительно одного класса);
- многозначная классификация (объект имеет признаки нескольких классов одновременно).

Задача выделения реперных горизонтов относится к бинарной классификации, а задача идентификации свит в скважине по каротажным данным – к многоклассовой. Для бинарной классификации выходным вектором является скалярное значение: 1 – репер, 0 – не репер. В задаче классификации каротажных данных по свитам определены 8 классов: Кияленская, Баженовская, Георгиевская, Куломзинская, Тарская, Васюганская, Тюменская, Палеозойские отложения. Существует два способа кодирования переменных номинального типа: Label Encoding и One Hot Encoding. Для задач классификации лучшим вариантом является выбор One Hot Encoding способа, в котором размер выходного вектора равен количеству классов, а i -ая компонента вектора указывает на вероятность наличия класса i у объекта и принимает значение в диапазоне, определяемым функцией активации нейронов выходного слоя.

Важным вопросом является определение того, что является входным вектором, что будет собственно подаваться на вход классификатора. Ряд работ основывается на поточечном составлении обучающей выборки, т.е. на вход сети подается n -мерный вектор признаков – значений n -различных каротажей для данной глубины скважины. Таким образом, чем больше будет значимых параметров в обучающем векторе, тем точнее будет обучаться классификатор. Однако, ввиду специфики каротажных данных, для ряда скважин с одного месторождения сложно составить репрезентативную выборку, т.к. мало участков с прописанными одинаковыми наборами каротажей. В данной работе

задача классификации, сегментации скважин на свиты и выделение реперов будет рассматриваться в качестве задачи анализа временных рядов.

Последовательность значений некоторой величины $X = \{x_1, x_2, \dots, x_n\}$, которая меняется во времени $\{t_1, t_2, \dots, t_n\}$ называется временным рядом. Задача классификации временных рядов состоит в поиске как можно более точного отображения $f: X \rightarrow \{1, 2, \dots, N\}$, которое принимает на вход временной ряд и ставит ему в соответствие один из N классов, а процесс поиска такого отображения называется обучением. Обучение осуществляется по обучающей выборке - набору $[(X_1, y_1), (X_2, y_2), \dots, (X_m, y_m)]$ объектов с известными ответами искомой функции. Результатом обучения будет являться функция, которая максимизирует качество классификации на тестовой выборке (X_i, y_i) размером K , т.е. вычисляется процент правильно классифицированных объектов: $Acc = \frac{1}{K} \sum_i^K [f(X_i) = y_i]$.

К распространенным задачам классификации временных последовательностей, например, относятся: распознавание речи, рукописного текста, жестов рук или головы, распознавание неисправностей (аварий) промышленных установок.

Для построения нейросетевого классификатора каротаж представляется как одномерный временной ряд, который разделяется («нарезается») на небольшие фрагменты, подпоследовательности, длиной в 40 – 80 м. (в зависимости от сети) с перекрытием в 50%. Отдельные фрагменты формируются так, что они не должны иметь пропущенных значений и соответствовать одному классу (свите).

4 Описание классификатора

Обобщенный алгоритм работы нейронной сети часто описывают в концепции «черного ящика», который переводит одни данные в другие, входной вектор – в выходное значение (Рисунок 4.1).

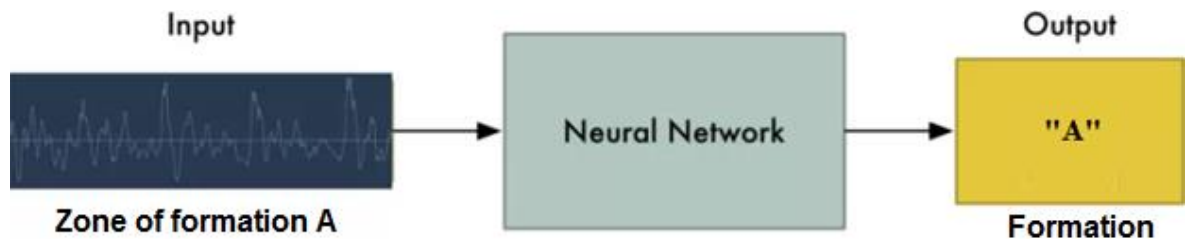


Рисунок 4.1 – Модель «черного ящика»

Известные методы классификации временных рядов за частую основываются на следующих методах:

- представление сигналов как точек в признаковых пространствах;
- структурное описание ряда;
- непосредственная работа по точкам сигнала.

Для решения задачи выделения свит в скважине по набору каротажных данных было решено построить каскад искусственных нейронных сетей прямого распространения, основываясь на идеи применения совокупности слабых классификаторов для достижения хорошей точности.

Конфигурация каждой сети была одинакова: полносвязная сеть прямого распространения с двумя скрытыми слоями по 50 и 30 нейронов, функция активации – гиперболический тангенс, на выходном слое из 8 нейронов применяется SoftMax (Рисунок 4.2).

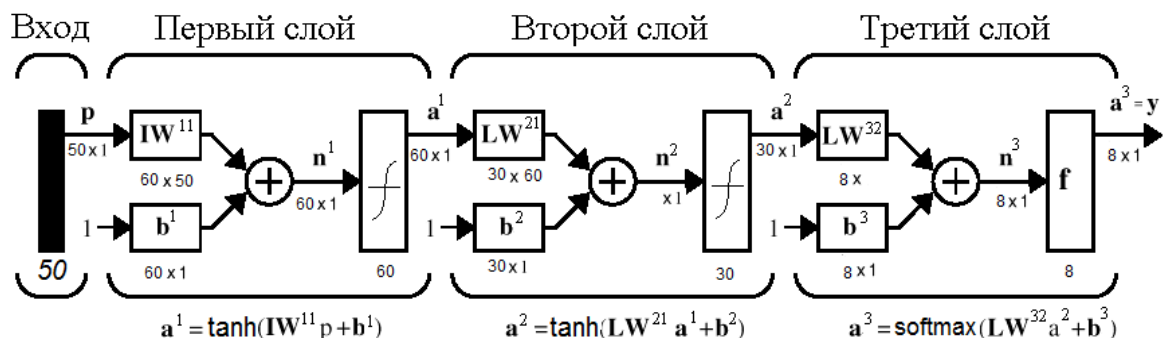


Рисунок 4.2 – Архитектура нейронной сети

На вход сети подается временной ряд соответствующего каротажа. Каскад состоит из сетей для каротажей гамма, собственной поляризации, индукционного, калипер, нейтронного гамма и акустического. Принадлежность входного ряда к классу определяется по взвешенной сумме голосов сетей.

Для решения задачи детектирования реперных горизонтов наилучших результатов удалось добиться с помощью применения на таких графических вероятностных моделях как скрытые Марковские модели (HMM).

Скрытая Марковская модель (Рисунок 4.3) — это модель из n скрытых состояний X и m наблюдаемых значений Y , которая определяется как $\lambda = (n, A, B, \pi)$, где A – матрица вероятностей переходов между состояниями, B – матрица вероятностей наблюдений выходных значений для каждого из состояний X , π – вектор вероятностей начальных состояний [16].

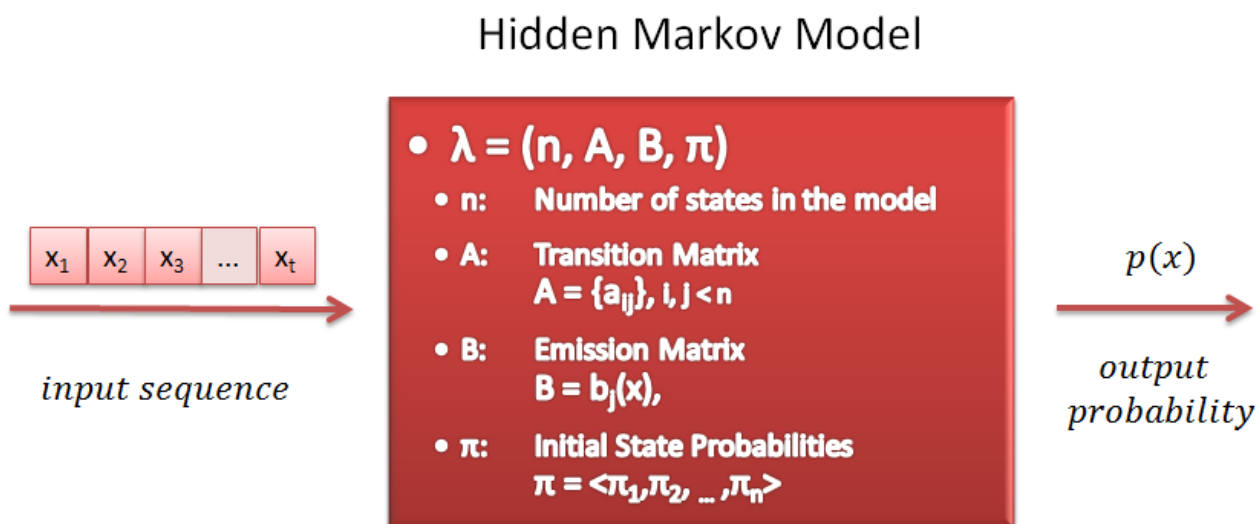


Рисунок 4.3 – Оценка вероятности последовательности с помощью HMM [17]

Для решения задачи классификации часто успешно применяются методы машинного обучения, основанные на вероятностных моделях, такие как скрытые Марковские модели и условные случайные поля со скрытыми состояниями (HCRF). HMM относится к классу порождающих моделей, а HCRF – к классу дискриминантных моделей. Обучение порождающих моделей состоит в нахождении параметров распределения вероятностей $p(x)$, наилучшим образом описывающих данные каждого класса, в то время как обучение дискриминантных моделей состоит в нахождении гиперплоскости,

наилучшим образом разделяющей данные каждого класса. Важным отличием является то, что дискриминантному классификатору необходимо p обучающих данных для достижения своей асимптотически минимальной ошибки классификации, тогда как порождающему классификатору необходимо $\log(p)$ обучающих данных [18]. Возможность обучения на малом количестве данных является важной особенностью классификаторов. К недостаткам порождающих моделей следует отнести то, что выбор формы функции плотности распределения вероятностей накладывает ограничения на применимость порождающих моделей, так как не все данные соответствуют тем или иным известным параметрическим семействам распределений вероятностей. Кроме того, так как наблюдаемые данные представляют собой многомерные вектора признаков, то оценка параметров распределений в случае наличия линейных зависимостей среди признаков вектора является затруднительной или невозможной [19].

Главным преимуществом реализации бинарного классификатора с помощью НММ в отличие от нейронных сетей является обучение на многомерных последовательностях различной длины. Вначале для наборов входных данных определяет набор k каротажных методов, присутствующий в каждой скважине. Затем по данным формируются обучающие k -мерные последовательности $X = \{x_1, x_2, \dots, x_t\}$, где $x_i = \{w_1, w_2, \dots, w_k\}$ – упорядоченные значения каротажей для точки глубины x_i . Обучающая выборка $O = \{X_1, X_2, \dots, X_l\}$ содержит k -мерные фрагменты данных скважины переменной длины t , описывающие реперные пласты. Для дополнения обучающей выборки отрицательными (отсутствие репера) примерами между каждыми двумя реперными горизонтами формируется последовательность случайной длины, описывающая пласт.

Общей характеристикой скрытых Марковских моделей является топология перехода состояний. Анализируя последовательности реперов и горных пород по глубине, при нахождении в определенном состоянии, нет возможности вернуться к состоянию, которое было ранее. Таким образом,

допускаются только переходы состояний вперед, выбирается прямая топология (Рисунок 4.4).

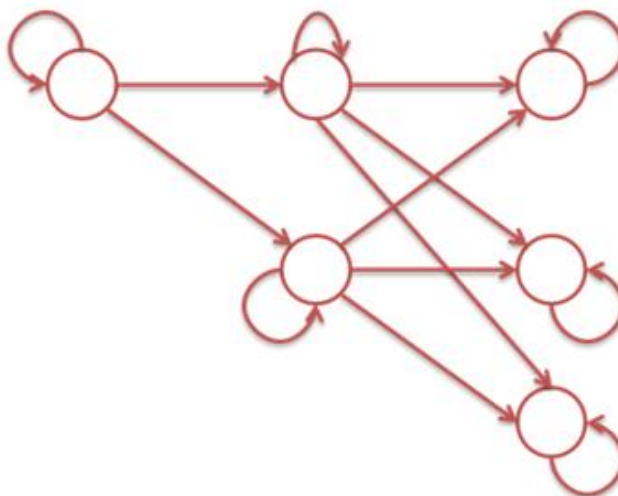


Рисунок 4.4 – Пример топологии состояний «только вперед»

Скрытая Марковская модель с $n=5$ скрытыми состояниями, k -мерным нормальным распределением вероятностей наблюдений выходных значений B , создается в ходе обучения на выборке O с помощью алгоритм Баума-Велша.

Для каждого типа последовательности (т. е. для репера и горной породы) создается НММ, которая выдает вероятности принадлежности к данному классу. На основе данных оценок строится классификатор максимального правдоподобия.

После создания бинарного классификатора реперов не менее важным вопросом является метод его применения на скважине для выделения реперов. Для этого используется принцип скользящего окна: окно заданного размера протаскивается по всей скважине и фрагмент каротажа, попавшего в окно, подается на вход классификатора. Центральной точке данного фрагмента присваивается полученный выход с классификатора (Рисунок 4.5).

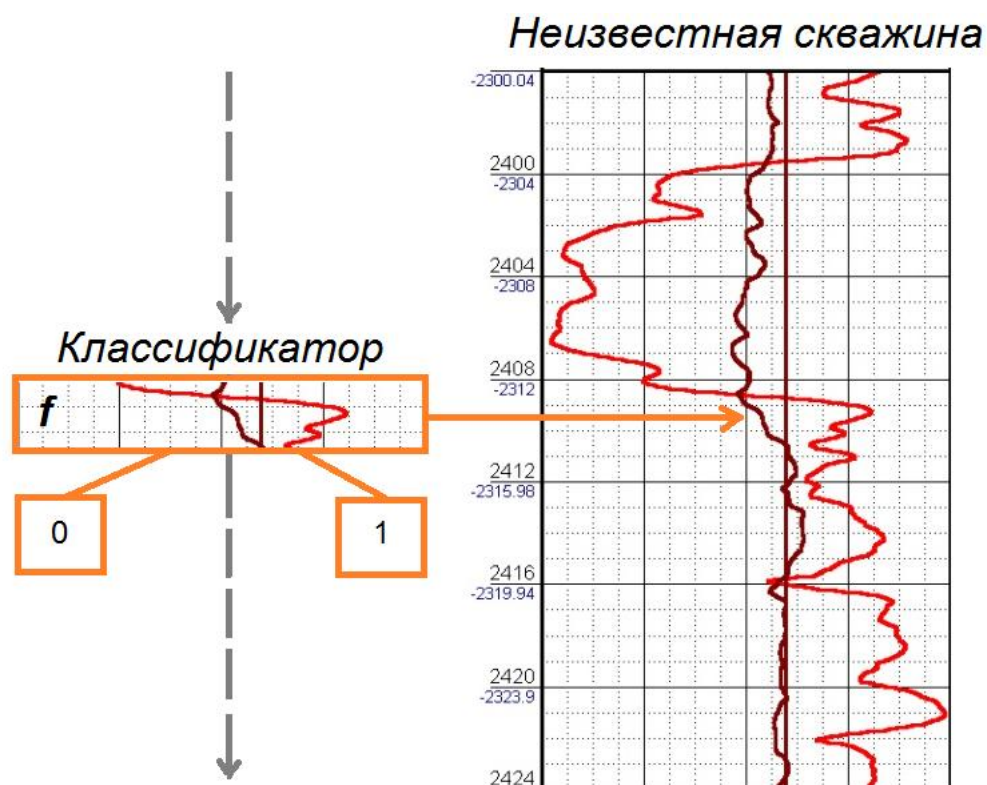


Рисунок 4.5 – Детектирование реперов скользящим окном

В данном подходе возможна ситуация, когда детектируются маркеры малой мощности (менее метра). Для этого осуществляется второй проход по размеченной скважине, при котором удаляются все выделенные реперы с мощностью, меньше заданной.

После нахождения реперных горизонтов в неизвестной скважине следует решать задачу корреляции соответствующих реперов из разных скважин. Для этого применяется алгоритм динамической трансформации временной шкалы (DTW-алгоритм, dynamic time warping). Данный алгоритм позволяет найти оптимальное соответствие между временными последовательностями разной длины, сдвинутыми или растянутыми относительно друг друга (Рисунок 4.6).

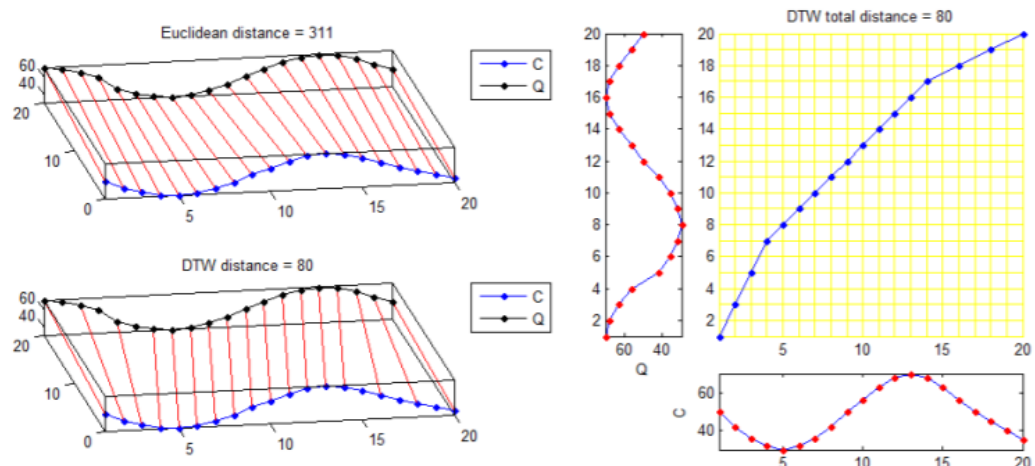


Рисунок 4.6 – Алгоритм динамической трансформации временной шкалы

Рассмотрим два временных ряда $Q = q_1, q_2, \dots, q_n$, и $C = c_1, c_2, \dots, c_m$. Вначале строится матрица расстояний D порядка $n \times m$, где $D_{i,j} = d(i, j)$.

Затем строится матрица трансформаций K порядка $n \times m$:

$$K_{i,j} = \begin{cases} D_{i,j} + \min(K_{i-1,j}, K_{i-1,j-1}, K_{i,j-1}), & \text{если } i > 1 \text{ и } j > 1 \\ D_{i,j}, & \text{если } i = 1 \text{ и } j = 1. \end{cases}$$

После заполнения матрицы деформации строится путь трансформации $p = (p_1, p_2, \dots, p_k)$ – последовательность элементов матрицы трансформации, которая минимизирует расстояние между траекториями, а также должен удовлетворять граничным условиям, условиям непрерывности и монотонности [20].

Среди всех возможных путей трансформации, которые удовлетворяют условиям, выбирается минимальный. DTW расстояние между двумя последовательностями через оптимальный путь трансформации выражается следующим образом: $D_{DTW}(Q, C) = \frac{p_k}{k}$.

Для автоматической корреляции реперных горизонтов между двумя скважинами для каждого репера первой вычисляется сумма DTW-расстояний по каждому признаку-каротажу репера для всех реперов второй скважины. Затем по минимуму суммы расстояний определяется коррелирующий репер. Для минимизации вычислений и улучшения точности корреляции проводится первоначальное выравнивание скважин по подошве Баженовской свиты, а затем автокорреляция только с соседними реперами в пределах 200 метров.

5 Анализ результатов исследования

В ходе выполнения данной магистерской диссертации была разработана программа, позволяющая загружать данные скважины с форматов .csv и .xml. Пользователь может вручную выделять реперные горизонты, на визуализированных каротажных диаграммах, конфигурировать классификатор и обучать на выбранных им скважинах. В результате программа выделяет реперные горизонты на неразмеченных скважинах. Пользователь может варьировать параметры классификации для достижения лучшего результата, а также самостоятельно редактировать результаты автоматического поиска.

5.1 Анализ результатов решения задачи классификации

Для решения задачи классификации горных пород по формациям использовались данные каротажных исследований, полученные при исследовании скважин месторождения С. Данный выбор обусловлен наличием данных об абсолютных отметках залегания исследуемых формаций, а также наибольшим количеством исходных данных для анализа.

В целом, как отмечали ряд исследователей в своих работах, чем больше различного рода каротажных исследований имеется, тем точнее будет решена задача классификации. Так в данном случае использовался следующий набор каротажных данных: ГК, ПС, НГК, АК, ИК, ДС. Все эти каротажи измеряют разные физические параметры, которые характеризуют массив горных пород. При использовании такого вида машинного обучения, как нейронные сети подавать наборы данных которые коррелируют между собой бессмысленно, так как это вызовет лишь увеличение времени обучения сети, без увеличения точности результатов. Именно поэтому не были добавлены данные полученные с электрических зондов различной глубинной, так как ИК измеряет проводимость горных пород, а проводимость – есть обратная величина сопротивления.

В главе анализ предметной области были приведены различные методы машинного обучения и виды нейронных сетей, применяемых для задачи классификации. При проведении исследования в ходе выполнения данной магистерской диссертации были рассмотрены 5 основных типов классификаторов: искусственная нейронная сеть, каскад сетей, байесовский классификатор, машины опорных векторов, а также метод k-ближайших соседей.

При решении задачи классификации выборка входных данных была разбита в пропорции 66% на обучающую выборку и 34% на тестовую. Весь объем входных данных был обработан в соответствии с методами, описанными в главе 4, данной магистерской диссертации.

Входные данные подразделялись на следующие классы: отложения кияленской свиты, тарской, куломзинской, баженовской, георгиевской, васюганской, тюменской, палеозойские отложения.

Использование искусственной нейронной сети как классификатора для исследуемых данных дало успешный результат, практически все формации имеют сходимость данных выше 70%. Рисунок 5.1 иллюстрирует матрицу точности произведенной классификации, с помощью нейронной сети. Центральная диагональ матрицы, показывает процент верности произведенной классификации. Наихудший результат был получен для отложений, приуроченных к: васюганской – 49.9%, тарской – 44.3%, а также георгиевской – 61.7% свит. Данный результат может быть объясним с позиции небольшого количества точек измерения. Наилучшие результаты были получены для отложений киялинской свиты – 94.7%, палеозойских отложений – 93%, тюменской свиты – 92.6% и баженовской свиты – 92.1%, а также отложений куломзинской свиты – 84.8%.

Для отложений васюганской и тарской свит можно заметить одну общую закономерность, наибольшее количество значений, отнесенных к другому классу, было отнесено к отложениям выше/нижележащей свиты. Так 34.2 % точек данных отложений васюганской свиты были ошибочно отнесены к

тюменской, а 43.5% отложений тарской свиты ошибочно классифицированы как отложения вышележащей киялинской свиты, а 11.2 % к нижележащей куломзинской свите.

		Predicted								Σ
		Bazhenovskaya	Georgievskaya	Kiyalinskaya	Kulomzinskaya	Paleozoic	Tarskaya	Tumenskaya	Vasuganskaya	
Actual	Bazhenovskaya	92.1 %	2.0 %	0.0 %	2.0 %	0.0 %	0.0 %	2.8 %	1.0 %	492
	Georgievskaya	10.6 %	61.7 %	0.0 %	20.2 %	0.0 %	0.0 %	1.1 %	6.4 %	94
	Kiyalinskaya	0.0 %	0.0 %	94.7 %	3.2 %	0.0 %	1.6 %	0.3 %	0.2 %	14058
	Kulomzinskaya	0.1 %	0.0 %	12.1 %	84.8 %	0.0 %	1.3 %	1.0 %	0.6 %	5080
	Paleozoic	0.0 %	0.0 %	0.0 %	0.0 %	93.0 %	0.0 %	6.9 %	0.1 %	3047
	Tarskaya	0.0 %	0.0 %	43.5 %	11.2 %	0.0 %	44.3 %	0.5 %	0.4 %	1405
	Tumenskaya	0.1 %	0.0 %	0.6 %	1.3 %	1.2 %	0.1 %	92.6 %	4.1 %	6665
	Vasuganskaya	0.7 %	0.7 %	7.6 %	5.7 %	0.2 %	0.9 %	34.2 %	49.9 %	1519
Σ		486	82	14699	5115	2918	935	7024	1101	32360

Рисунок 5.1 – Матрица точности классификации нейронной сети.

Следующим видом классификатора является каскад нейронных сетей, стоит особо отметить быстроту обучения данного классификатора, которая является наименьшей из всех используемых в рамках данного проекта.

Точность классификации для каскада сетей представлена на Рисунке 5.2. Так, по сравнению с предыдущим классификатором, точность классификации возросла практически для всех групп отложений. Кроме отложений киялинской свиты, для которой точность снизилась на 2.9%, однако при этом составила 91.8%. Точность классификации отложений васюганской, тарской и георгиевской свит составила 68.3%, 55.4% и 70.2% соответственно.

		Predicted								Σ
		Bazhenovskaya	Georgievskaya	Kiyalinskaya	Kulomzinskaya	Paleozoic	Tarskaya	Tumenskaya	Vasuganskaya	
Actual	Bazhenovskaya	94.7 %	1.2 %	0.0 %	1.4 %	0.0 %	0.0 %	2.2 %	0.4 %	492
	Georgievskaya	8.5 %	70.2 %	0.0 %	7.4 %	0.0 %	0.0 %	7.4 %	6.4 %	94
	Kiyalinskaya	0.0 %	0.0 %	91.8 %	3.7 %	0.0 %	3.8 %	0.3 %	0.5 %	14058
	Kulomzinskaya	0.1 %	0.1 %	9.5 %	85.9 %	0.0 %	2.3 %	1.1 %	1.0 %	5080
	Paleozoic	0.0 %	0.0 %	0.0 %	0.0 %	97.8 %	0.0 %	2.2 %	0.0 %	3047
	Tarskaya	0.0 %	0.0 %	35.0 %	8.3 %	0.0 %	55.4 %	0.8 %	0.6 %	1405
	Tumenskaya	0.1 %	0.2 %	0.7 %	0.9 %	1.1 %	0.1 %	91.7 %	5.2 %	6665
	Vasuganskaya	0.2 %	0.5 %	4.0 %	2.8 %	0.3 %	1.5 %	22.3 %	68.3 %	1519
Σ		485	95	13985	5126	3060	1452	6637	1520	32360

Рисунок 5.2 – Матрица точности классификации каскада сетей

Если принять пороговое значение точность в 70%, как меру верности произведенной классификации, то отложения лишь двух свит можно считать классифицированными неверно. А в целом васюганская свита почти достигает установленного порогового значения для задачи классификации.

Следующим анализируемым методом является байесовский классификатор. Данный классификатор также имел высокую скорость обучения наравне с каскадом сетей. На Рисунке 5.3 иллюстрируется полученная точность результатов при использовании байесовского классификатора для решения задачи классификации.

		Predicted								Σ
		Bazhenovskaya	Georgievskaya	Kiyalinskaya	Kulomzinskaya	Paleozoic	Tarskaya	Tumenskaya	Vasuganskaya	
Actual	Bazhenovskaya	75.8 %	0.0 %	0.0 %	3.9 %	0.0 %	0.0 %	20.1 %	0.2 %	492
	Georgievskaya	44.7 %	27.7 %	0.0 %	24.5 %	0.0 %	0.0 %	2.1 %	1.1 %	94
	Kiyalinskaya	0.0 %	0.0 %	79.9 %	10.8 %	0.9 %	1.0 %	6.9 %	0.5 %	14058
	Kulomzinskaya	0.3 %	1.7 %	24.6 %	52.5 %	0.1 %	0.7 %	16.5 %	3.5 %	5080
	Paleozoic	0.0 %	0.0 %	0.0 %	0.0 %	85.1 %	0.0 %	14.8 %	0.1 %	3047
	Tarskaya	0.0 %	0.0 %	57.9 %	10.0 %	5.8 %	12.5 %	10.0 %	3.8 %	1405
	Tumenskaya	0.7 %	0.1 %	0.0 %	2.2 %	6.5 %	0.0 %	82.1 %	8.4 %	6665
	Vasuganskaya	3.3 %	0.3 %	1.1 %	8.8 %	6.1 %	0.3 %	36.1 %	44.2 %	1519
Σ		530	121	13313	4648	3331	363	8520	1534	32360

Рисунок 5.3– Матрица точности байессова классификатора

Так с достаточным уровнем точности байесовский классификатор может классифицировать лишь отложения четырех свит: баженовская, киялинская, тюменская и отложения палеозоя. Наихудший результат получен для отложений тарской свиты 12.5%, а также георгиевской свиты 27.7%. Таким образом, байесовский классификатор не обеспечивает точное определение половины типов анализируемых отложений.

Машины опорных векторов, используемые для решения задачи классификации показали наихудший результат из всех методов. Точность решения задачи классификации с применением метода машин опорных векторов приведена на Рисунке 5.4. Уверенно классифицируются лишь 2 из 8 анализируемых типов отложений. Так наилучшие результат получены для

отложений баженовской свиты и палеозойских отложений, 89.6% и 93.8% соответственно. Наихудшие результаты, менее 20% точности, были получены для отложений куломзинской, тарской и тюменской свит. При этом стоит особенно отметить, что метод имел скорость обучения ниже, чем приведенные ранее классификаторы.

		Predicted								Σ
		Bazhenovskaya	Georgievskaya	Kiyalinskaya	Kulomzinskaya	Paleozoic	Tarskaya	Tumenskaya	Vasuganskaya	
Actual	Bazhenovskaya	89.6 %	3.0 %	0.0 %	3.5 %	0.0 %	0.2 %	0.8 %	2.8 %	492
	Georgievskaya	25.5 %	48.9 %	0.0 %	12.8 %	0.0 %	0.0 %	1.1 %	11.7 %	94
	Kiyalinskaya	0.0 %	0.0 %	33.7 %	2.3 %	0.0 %	12.5 %	7.4 %	44.1 %	14058
	Kulomzinskaya	6.5 %	1.5 %	9.1 %	3.3 %	0.3 %	9.0 %	14.0 %	56.4 %	5080
	Paleozoic	0.0 %	0.0 %	2.4 %	1.1 %	93.8 %	0.1 %	1.7 %	1.0 %	3047
	Tarskaya	0.0 %	0.0 %	17.0 %	0.3 %	0.1 %	15.5 %	10.2 %	56.9 %	1405
	Tumenskaya	0.2 %	0.0 %	8.7 %	2.1 %	28.5 %	3.8 %	13.1 %	43.5 %	6665
	Vasuganskaya	1.6 %	1.3 %	16.5 %	2.9 %	4.9 %	5.7 %	7.4 %	59.8 %	1519
Σ		840	159	6338	745	4849	2772	2936	13721	32360

Рисунок 5.4 – Матрица точности для метода машин опорных векторов

Дополнительно к четырем описанным выше классификаторам использовался классификатор, основанный на применении метода *k*-ближайших соседей. Матрица точность данного метода классификации приведена на Рисунке 5.5. Так низкая точность классификации выявлена для отложений васюганской, тарской и георгиевской свит: 45.3%, 24.1% и 55.3% соответственно. Высокая степень точности классификации для трех групп: отложения палеозоя 90.8%, баженовской свиты 92.9% и киялинской свиты 94.1%.

		Predicted								Σ
		Bazhenovskaya	Georgievskaya	Kiyalinskaya	Kulomzinskaya	Paleozoic	Tarskaya	Tumenskaya	Vasuganskaya	
Actual	Bazhenovskaya	92.9 %	0.2 %	0.0 %	1.4 %	1.0 %	0.0 %	4.1 %	0.4 %	492
	Georgievskaya	4.3 %	55.3 %	2.1 %	6.4 %	0.0 %	0.0 %	28.7 %	3.2 %	94
	Kiyalinskaya	0.0 %	0.0 %	94.1 %	3.7 %	0.0 %	1.7 %	0.3 %	0.1 %	14058
	Kulomzinskaya	0.0 %	0.0 %	17.6 %	80.3 %	0.1 %	0.5 %	1.3 %	0.3 %	5080
	Paleozoic	0.1 %	0.0 %	0.0 %	0.4 %	90.8 %	0.0 %	8.4 %	0.4 %	3047
	Tarskaya	0.0 %	0.0 %	67.1 %	7.7 %	0.0 %	24.1 %	0.4 %	0.7 %	1405
	Tumenskaya	0.3 %	0.2 %	2.1 %	2.9 %	2.9 %	0.1 %	87.5 %	4.2 %	6665
	Vasuganskaya	0.3 %	0.2 %	7.3 %	6.7 %	2.3 %	1.1 %	36.8 %	45.3 %	1519
Σ		491	68	15313	5026	3002	632	6804	1024	32360

Рисунок 5.5 – Матрица точности для метода *k*-ближайших соседей

Сравнительный анализ используемых методов классификации приведен в Таблице 5.1.

Таблица 5.1 – Сравнение используемых классификаторов

Название классификатора	Скорость обучения	Общая точность результатов; %
Искусственная нейронная сеть	Средняя	88.2
Каскад сетей	Высокая	88.7
Баевсов классификатор	Высокая	71.7
Машины опорных векторов	Низкая	31.6
k-ближайших соседей	Средняя	84.8

Совокупный анализ используемых методов классификации однозначно свидетельствует на преимущество использования каскада сетей. Так как он обеспечивает высокую степень точности и обладает наивысшей скоростью обучения. А также способен классифицировать практически все типы отложений с убедительной точностью. Наихудшим классификатором является метод машин опорных векторов, так как он имеет наименьшую точность и требует наибольшее время для обучения.

Исходя из проделанной на данном этапе работы, можно сделать вывод, что отложения двух типов классифицируются с наивысшей точностью, при применении любых методов классификации– это отложения бажендовской свиты и палеозойские отложения. Таким образом, данные отложения могут быть использованы для дальнейшей цели выделения реперных горизонтов.

5.2 Анализ результатов выделения реперных горизонтов

Для решения задачи выделения реперных горизонтов использовались данные полученные с месторождения D Томской области, которое также располагается в пределах Нюрольской впадины. Данное месторождение было выбрано по причине наличия экспертной схемы корреляции, которая будет использована для сравнения полученных результатов.

На месторождении D экспертами выделены следующие реперные горизонты: баженовская свита, сложенная битуминозными аргиллитами, обладающими высокими значениями ГК; 8 пластов, том числе регионально протяженный угольный пласт У1, который также считается региональным реперным горизонтом. Экспертная схема выделения реперных горизонтов и корреляция скважин представлена на Рисунке 5.6.

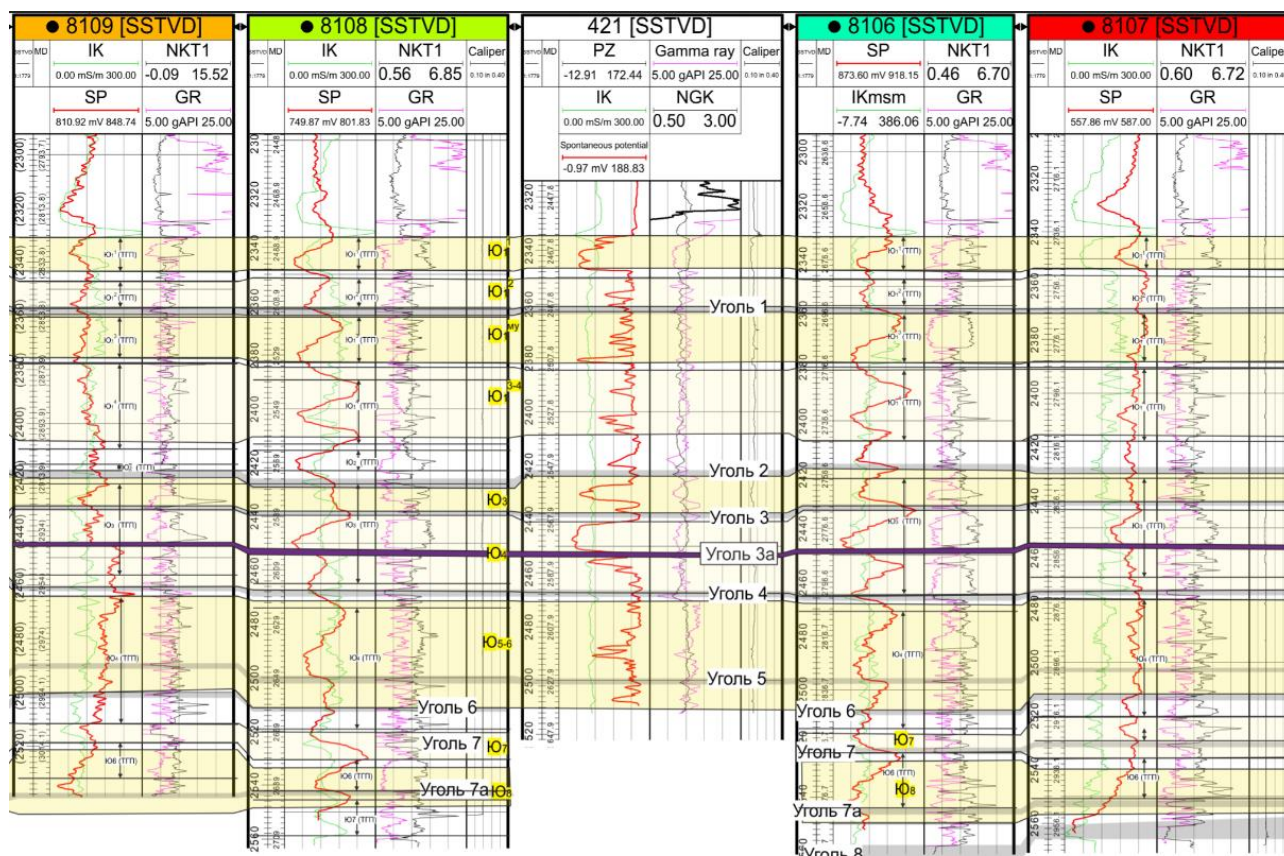


Рисунок 5.6 – Экспертная схема корреляции для месторождения D

Для проверки правильности работы самостоятельно был добавлен еще один реперный горизонт: отложения георгиевской свиты.

Обучение классификатора происходило на основе выделенных реперов в 4-х скважинах, данные 5-й скважины являются тестовыми. В результате классификатору удалось выделить все основные реперные горизонты, выделенные экспертами, а также выделить отложения георгиевской свиты (Рисунок 5.7).

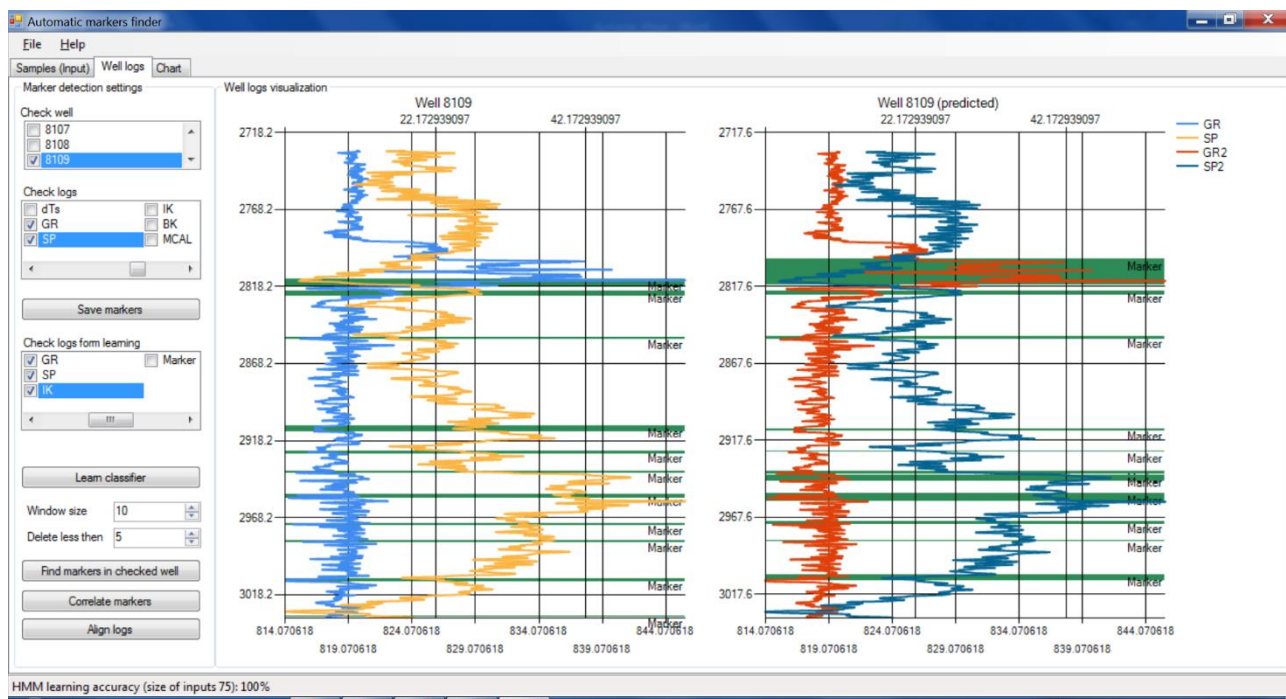


Рисунок 5.7 – Результат выделения реперов классификатором на скв. №8109 (слева определено экспертами, справа определено классификатором)

Стоит отметить качество выделения региональных реперов, таких как: баженовская свита и угольный пласт У1, а также отложения георгиевской свиты. При этом достаточно было выделить малую часть баженовской свиты как репер для обучения, чтобы классификатор выделил ее полностью. Стоит дополнительно указать, что из 11 указанных реперных горизонтов, не выделенным оказался самый маломощный угольный пласт, выделенный экспертами. Эту проблемы можно решить за счет ручной регулировки параметров обработки классифицируемой последовательности: размер окна прохода по каротажным данным, и минимальная мощность выделяемого пласта в м. Однако при слишком большом уменьшении данных параметров классификатор обнаружит маломощные слои, которые не были учтены экспертами. Так на Рисунке 5.8 представлены результаты, полученные при ручной регулировке параметров обработки классифицируемой последовательности, при которой наименьший угольный реперный пласт был выделен.

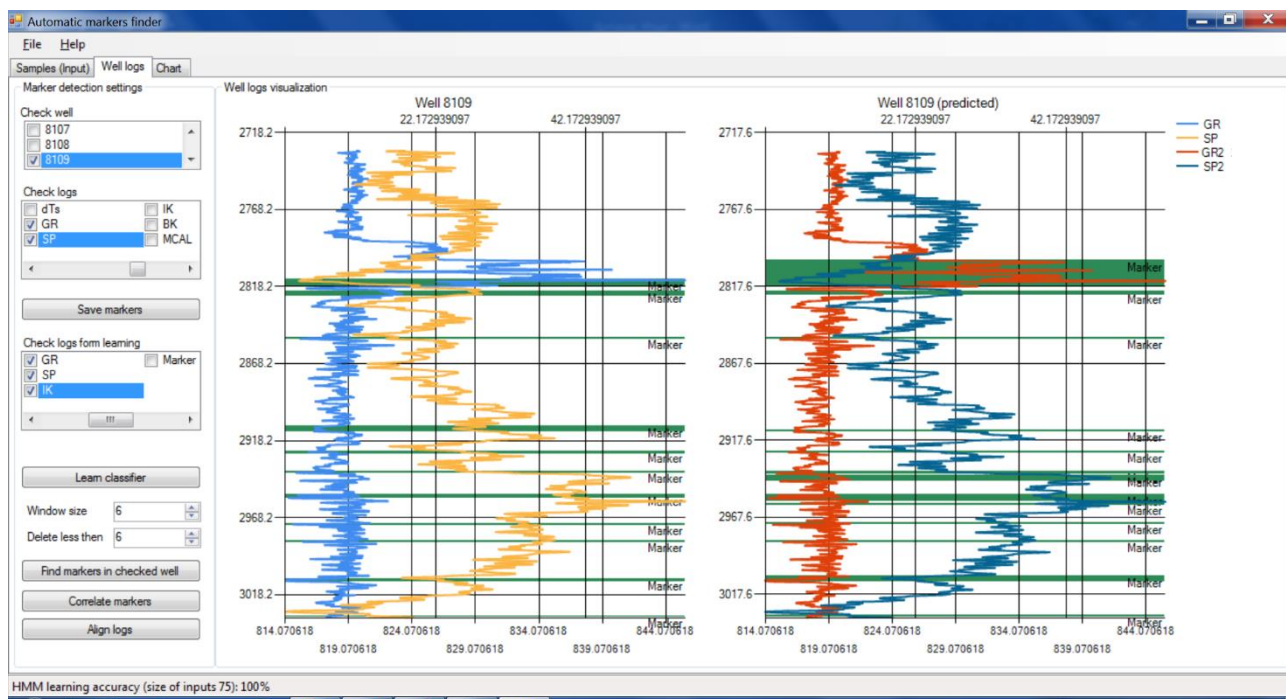


Рисунок 5.8– Результаты выделения реперов классификатором на скв. №8109 при ручной регулировке параметров обработки классифицируемой последовательности

6 Финансовый менеджмент, ресурсоэффективность и ресурсосбережение

Целью данного раздела магистерской диссертации является анализ перспективности научно-исследовательского проекта. В рамках данного проекта производилась работа по созданию программного обеспечения для обеспечения проведения геолого-геофизической типизации разреза. Анализ себестоимости данного программного обеспечения, а также основные финансовые расчеты представлены в данном разделе магистерской диссертации.

Необходимым условием является анализ ресурсоэффективности проекта, необходимой материальной базы. Также в части менеджмента проекта важнейшей частью является планирование работы над проектом и обозначение целей, обозначающих вектор развития проекта и методов, которыми данные цели предполагается достичь.

6.1 Анализ преимуществ реализации проекта

Целевым рынком для программы автоматической геолого-геофизической типизации разреза является рынок программного обеспечения для обработки данных геофизических исследований. В данной магистерской диссертации рассматривается использование нейронных сетей, для решения задачи типизации геолого-геофизического разреза. Данный аспект является уникальным для такого типа программного обеспечения, что делает его уникальным относительно других.

Возможные варианты исполнения задачи геолого-геофизической интерпретации представлены в Таблице 6.1 .

Таблица 6.1 – Варианты решения задачи геолого-геофизической типизации разреза

Критерии оценки	Способ решения		
	Ручной способ решения	Программные комплексы автоматической типизации	Программные комплексы автоматической типизации с внесением экспертных критериев
Скорость	низкая	высокая	высокая
Точность	средняя	высокая	высокая
Геологическая верность корреляции	высокая	средняя	высокая

Согласно анализу вариантов решения задачи типизации геолого-геофизической типизации разреза, применение программных комплексов с внесением экспертных критериев является наилучшим. Решение принятое в пользу такого варианта реализации обуславливает возможность внесение комплексного понимания обстановки седиментации экспертом в процесс типизации, что в свою очередь должно приводить к увеличению верность производимой корреляции.

В качестве компаний-потребителей выступают как подрядные геофизические организации и проектные институты, так и компании операторы, которые имеют собственную геофизическую службу.

6.2 Планирование управления научно-техническим проектом

6.2.1 План проекта

Реализация любого программного продукта предусматривает глубокий анализ и четкое планирование занятости. Главнейшей целью данного этапа является рационализация использования ресурсов, которыми располагает, выполняющий данный проект субъект. При правильной организации трудового процесса происходит максимизация творческого и рабочего потенциала и как следствие сокращается время реализации проекта, при сохранении или улучшении качества производимого продукта. Данный раздел представляет полный перечень выполненных работ. Для каждого из этапов определены их

исполнители и ответственные лица, а также рациональная продолжительность периода работы. Наиболее наглядным итогом запланированных работ является сетевой график реализации проекта либо линейный вариант. В виду того, что при реализации данной магистерской диссертации исполнителей двое: НР – научный руководитель и ВМ – выпускник-магистр, то наиболее информативным представляется линейный график реализации проекта. Хронология выполнения работ упорядочена в Таблице 6.2.

Таблица 6.2 – Календарный план проекта

Код работы	Название	Длительность, дни	Дата начала работ	Дата окончания работ	Состав участников
1	Составление и утверждение технического задания	4	27.06.18	01.07.18	НР
2	Подбор и изучение материалов	12	01.07.18	13.07.18	ВМ
3	Календарное планирование работ	2	13.07.18	15.07.18	НР, ВМ
4	Изучение основ программирования	14	15.07.18	29.07.18	ВМ
5	Изучение и анализ применяемых методов классификации данных	7	29.07.18	01.08.18	ВМ
6	Предварительное обсуждение результатов выполненных работ. Анализ вариантов оптимизации проекта	1	01.08.18	02.08.18	ВМ, НР
7	Анализ применения нейронной сети для решения мелкомасштабной задачи.	4	02.08.18	06.08.18	ВМ, НР
8	Реализация двух анализируемых задач в программном ПО	4	06.08.18	10.08.18	ВМ
9	Тестирование и обсуждение произведенной работы	1	10.08.18	11.08.18	ВМ, НР
10	Выполнение раздела финансовый менеджмент	12	01.08.18	13.08.18	ВМ
11	Выполнение раздела оценки безопасности и экологичности проекта	2	13.08.18	15.08.18	ВМ
12	Анализ выполненной работы и оформление диссертации	2	15.08.18	17.08.18	ВМ, НР
Итого:		55	27.06.18	17.08.18	

Диаграмма Ганта представляет собой столбчатую диаграмму, которая отражает стадии проекта и иллюстрирует основные календарные этапы его развития. На Рисунке 6.1 представлена диаграмма Ганта иллюстрирующая план-график выполнения работ по магистерской диссертации по типизация геолого-геофизических разрезов.

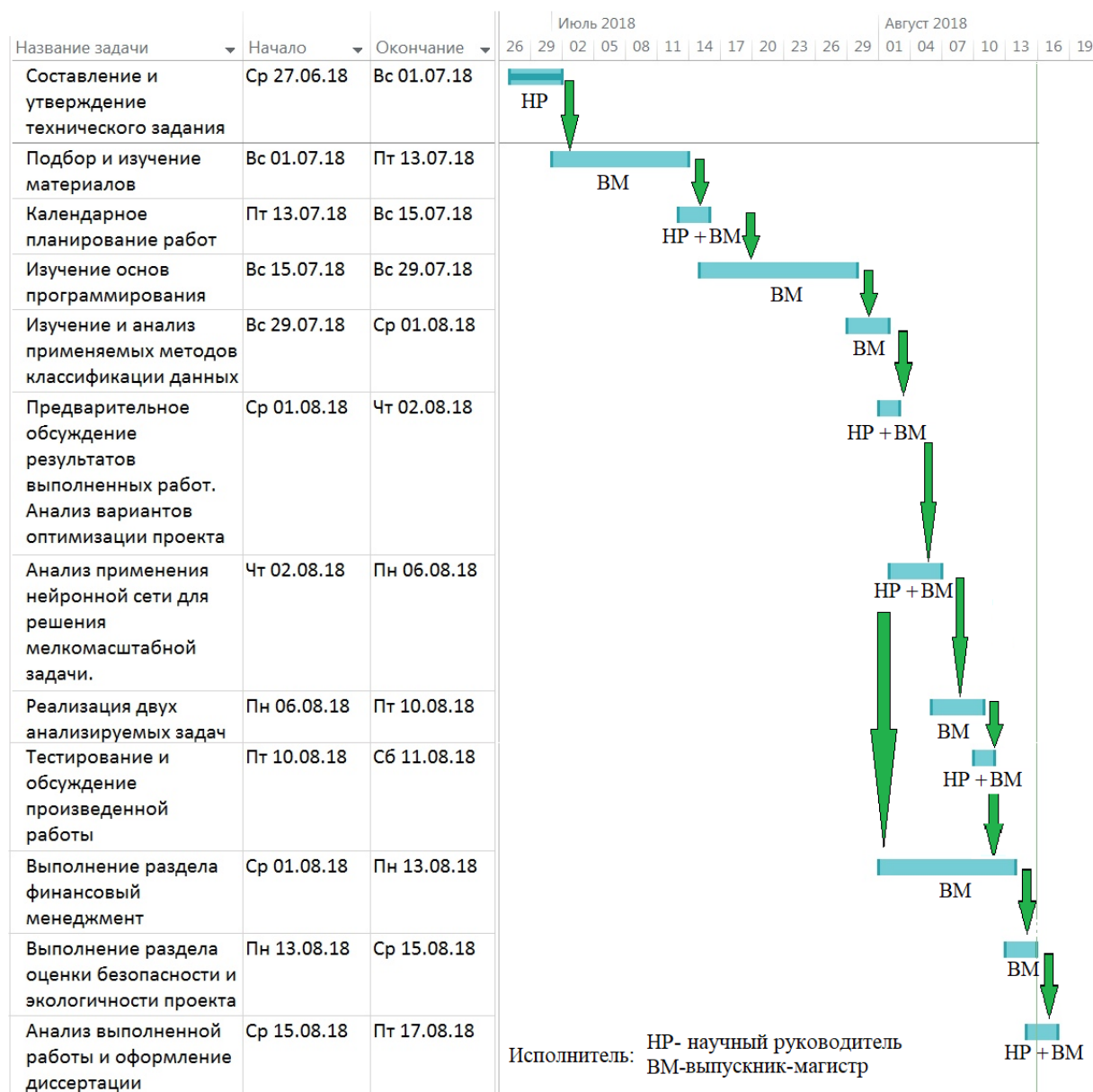


Рисунок 6.1 – Диаграмма Ганта выполнения магистерской диссертации

6.2.2 Определение трудоёмкости выполнения работ

Основную часть себестоимость программного обеспечения, создаваемого для типизации геолого-геофизических разрезов, составляют

трудовые затраты. Трудоёмкость определялась с помощью метода экспертной оценки и рассчитывалась в человеко-днях. При этом, для расчета ожидаемого значения трудоёмкости $t_{ож\ i}$ использовалась следующая формула:

$$t_{ож\ i} = \frac{3t_{min\ i} + 2t_{max\ i}}{5}, \quad (1)$$

где $t_{ож\ i}$ – ожидаемая трудоемкость выполнения i -ой работы, чел.-дн.; $t_{min\ i}$ и $t_{max\ i}$ – минимально и максимально возможная трудоемкость работ выполнения i -ой работы соответственно, чел-дн.

По рассчитанной ожидаемой трудоёмкости работы определяется ее продолжительность в рабочих днях $t_{раб\ i}$. Учетом параллельность выполнения некоторых работ сразу несколькими исполнителями:

$$t_{раб\ i} = \frac{t_{ож\ i}}{K_{вн}} \cdot K_{д}, \quad (2)$$

где $K_{вн}$ – коэффициент выполнения работ ($K_{вн} = 1,1$); $K_{д}$ – коэффициент, учитывает время необходимое для согласования работ ($K_{д} = 1,2$).

Для удобства расчета графика выполнения работ и их длительности необходим перевод из рабочих дней в календарные дни:

$$T_{КД\ i} = t_{раб\ i} \cdot T_{К}, \quad (3)$$

где $T_{КД\ i}$ и $t_{раб\ i}$ – продолжительность выполнения i -ой работы в календарных днях (дн.) и в рабочих днях (раб. дн.) соответственно; где $T_{К}$ – коэффициент календарности.

Коэффициент календарности рассчитывается по формуле:

$$T_{К} = \frac{T_{КАЛ}}{T_{КАЛ} - T_{ВД} - T_{ПД}}, \quad (4)$$

где $T_{КАЛ}$ – календарные дни ($T_{КАЛ} = 365$); $T_{ВД}$ – выходные дни; $T_{ПД}$ – праздничные дни ($T_{ВД} + T_{ПД} = 118$). Коэффициент календарности для выполняемого научного исследования в 2018 году: $T_{К} = \frac{365}{365 - 118} \approx 1,478$.

Рассчитанные значения, полученные в календарных днях, для каждого вида работ в рамках магистерской диссертации $T_{КД\ i}$ следует округлять до целого числа. Расчетные временные показатели, полученные для выполнения магистерской диссертации, представлены в Таблице 6.3.

Таблица 6.3 – Трудозатраты на выполнение проекта

Наименование работ	Исполнители	Продолжительность работ в днях, дн.						Длительность работ в рабочих днях, раб. дн.		Длительность работ в календарных днях, дн.	
		$t_{min i}$		$t_{max i}$		$t_{ож i}$		$t_{раб i}$		$T_{кд i}$	
		В М	НР	ВМ	НР	ВМ	НР	ВМ	НР	ВМ	НР
Составление и утверждение ТЗ	НР		4		5		4.4		4.8		7.1
Подбор и изучение материалов	ВМ	12		13		12.4		13.5		20	
Календарное планирование работ	НР, ВМ	1	1	1	2	1	1.4	1.1	1.5	1.6	2.3
Изучение основ программирования	ВМ	14		20		16.4		17.9		26.4	
Изучение и анализ методов классификации	ВМ	7		10		8.2		8.9		13.2	
Предварительное обсуждение результатов работ. Анализ вариантов оптимизации проекта	ВМ, НР	0,5	0,5	1	1	0.7	0.7	0.8	0.8	1.1	1.1
Анализ применения нейронной сети для решения мелкомасштабной задачи	ВМ, НР	2	2	4	3	2.8	2.4	3.1	2.6	4.5	3.9
Реализация двух анализируемых задач в программном ПО	ВМ	4		6		4.8		5.2		7.7	
Тестирование и обсуждение работы	ВМ, НР	0.5	0.5	0.5	1	0.5	0.7	0.5	0.8	0.8	1.1
Выполнение раздела финансовый менеджмент	ВМ	12		15		13.2		14.4		21.3	
Выполнение раздела социальной ответственность	ВМ	2		4		2.8		3.1		4.5	
Анализ выполненной работы и оформление диссертации	ВМ, НР	1	1	2	2	1.4	1.4	1.5	1.5	2.3	2.3
Итого:						64.2	11	70	12	103.5	17.7

6.2.3 Бюджет научного исследования

Затраты на создание проекта – совокупная стоимость ресурсов, необходимых для создания проекта. Основной расчет стоимости выполнения магистерской диссертации по теме типизация геолого-геофизических разрезов с применением методов машинного обучения производился по следующим статьям расходов:

- заработная плата сотрудников, выполняющих проект;
- обязательные отчисления в фонды;
- подоходный налог;
- амортизационные отчисления;
- материальные затраты на создание проекта;
- прочие расходы.

6.2.3.1 Расчет заработной платы

Данная статья состоит из основной $Z_{\text{осн}}$ и дополнительной заработной платы $Z_{\text{доп}}$ научного руководителя и магистранта:

$$Z_{\text{зп}} = Z_{\text{осн}} + Z_{\text{доп}}.$$

Установлено, что размер дополнительной заработной платы должен составлять 15% от основной части. При этом, калькуляция основной части заработной платы выполняется с учетом трудоемкости выполняемого этапа и прямо пропорционален величине ежемесячного оклада исполнителя:

$$Z_{\text{осн}} = Z_{\text{м}} \cdot \frac{M}{F_{\text{д}}} \cdot T_{\text{раб}}, \quad (5)$$

где $Z_{\text{м}}$ – месячный оклад, руб.;

M – количество месяцев работы без отпуска в течение года ($M = 10,4$);

$F_{\text{д}}$ – годовой фонд рабочего времени ($F_{\text{д}} = 247$ раб. дн.).

Месячный должностной оклад работника:

$$Z_{\text{м}} = Z_{\text{б}} \cdot (k_{\text{пр}} + k_{\text{д}}) \cdot k_{\text{р}}, \quad (6)$$

где $Z_{\text{б}}$ – базовый оклад работника, руб.;

$k_{\text{пр}}$ – премиальный коэффициент ($k_{\text{пр}} = 1,15$), премия 15%;

$k_{\text{д}}$ – коэффициент доплат и надбавок ($k_{\text{д}} = 0,2$);

$k_{\text{р}}$ – районный коэффициент ($k_{\text{р}} = 1,3$).

Оклад выпускника магистра примем равным 15000 рублей в месяц. Проведем расчет аналогично расчету для научного руководителя. Выполненная калькуляция заработной платы исполнителей проекта приведена в Таблице 6.4.

Таблица 6.4 – Затраты на заработную плату

Исполнитель	Оклад, руб./мес.	Среднедневная ставка, руб./раб. дн.	Затраты времени, раб. дн.	Коэффициент, $1,15 \cdot (k_{\text{пр}} + k_{\text{д}}) \cdot k_{\text{р}}$	ФЗП, руб.
Научный руководитель	26300	1252.38	12	2.02	30357.69
Выпускник-магистр	15000	714.29	70	2.02	101000.61
Итого:					131358.3

6.2.3.2 Размер обязательных отчислений в фонды

Согласно законодательству РФ, ежемесячно из фонда заработной платы работников удерживается часть денежных средств, распределяемая в социальные фонды. Отчисление выплат в фонды является обязанностью работодателя. Отчисления из фонда заработной платы поступают в следующие фонды: фонд государственного социального страхования (ФГСС); пенсионный фонд РФ (ПФРФ); федеральный фонд обязательного медицинского страхования (ФФОМС). Совокупная величина отчислений во все внебюджетные фонды на сегодняшний момент составляет $k_{\text{внеб}} = 31,8\%$.

Произведем калькуляцию отчислений во внебюджетные фонды для научного руководителя и выпускника-магистра:

$$\text{НР: } Z_{\text{внеб}} = k_{\text{внеб}} \cdot (Z_{\text{осн}} + Z_{\text{доп}}) = 0.318 \cdot 30357,69 = 9653.75 \text{ руб.},$$

$$\text{ВМ: } Z_{\text{внеб}} = k_{\text{внеб}} \cdot (Z_{\text{осн}} + Z_{\text{доп}}) = 0.318 \cdot 101000.61 = 32118.19 \text{ руб.}$$

В результате произведенной калькуляции, сумма отчислений во внебюджетные фонды составила 41771.94 рубля.

6.2.3.3 Расчет подоходного налога

Со всех работающих граждан РФ взимается подоходный налог, ставка которого определяется в соответствии с законодательством РФ и в данном случае равна 13% от фонда заработной платы. Данный платеж отчисляется непосредственно в бюджет РФ.

Произведем расчет подоходного налога для научного руководителя и выпускника-магистра:

$$\text{НР: Нпд} = k_{\text{под.нал}} \cdot (З_{\text{осн}} + З_{\text{доп}}) = 0.13 \cdot 30357,69 = 3946.50 \text{ руб.},$$

$$\text{ВМ: Нпд} = k_{\text{под.нал}} \cdot (З_{\text{осн}} + З_{\text{доп}}) = 0.13 \cdot 101000.61 = 13130.08 \text{ руб.}$$

Где $k_{\text{под.нал}}$ – ставка подоходного налога

6.2.3.4 Расчет затрат на амортизацию специального оборудования

В данном разделе производится расчет амортизационных отчислений. В данном случае оборудованием на которое необходимо рассчитать амортизацию будет персональный компьютер и оргтехника используемая при выполнении проекта. Ориентировочная исходная стоимость оборудования указана в Таблице 6.5.

Таблица 6.5 – Амортизационная база

Наименование	Начальная стоимость, руб.
Персональный компьютер	60000
Оргтехника	80000

Расчет амортизационных отчислений производился по следующей формуле:

$$З_{\text{ам}} = \frac{H_A \cdot Ц_{\text{об}} \cdot t_{\text{рф}}}{F_d}, \quad (7)$$

где H_A – норма амортизации единицы оборудования (20%);

$Ц_{\text{об}}$ – балансовая стоимость оборудования;

F_d – действительный годовой фонд времени работы соответствующего оборудования (560 часов);

$t_{\text{рф}}$ – фактическое время работы оборудования;

Амортизационные взносы:

Персональный компьютер: $Z_{\text{АМ}} = 0,2 \cdot 60000 \cdot 560 \cdot \frac{1}{2075} = 3238,55 \text{ руб.}$,

Оргтехника: $Z_{\text{АМ}} = 0,2 \cdot 80000 \cdot 560 \cdot \frac{1}{2075} = 4318,07 \text{ руб.}$

Предположим совокупное время использования оргтехники равным, времени использования персонального компьютера. Таким образом суммарные амортизационные отчисления составили 7556.62 рубля.

6.2.3.5 Расчет материальных затрат

В материальные затраты на выполнение магистерской диссертации включены, затраты, связанные с процессом непосредственного выполнения диссертации. Составление отчета, расчеты и пометки, исправления записки все это включено в затраты. Таким образом затраты на канцелярские товары, носители данных: флэш-накопитель и компакт-диск, и прочие расходы. Расчетными данными материальных затрат являются фактические чеки о покупке исходных материалов. Прочие расходы рассчитаны как 15% от суммы остальных расходов. В Таблице 6.6 приведены материальные затраты проекта.

Таблица 6.6 – Материальные затраты при реализации проекта

Наименование	Количество	Цена, руб.	Затраты, руб.
Ручка шариковая	2	120	240
Упаковка бумаги А4	1	350	350
Карандаш НВ	2	25	50
Накопители данных	1	750	750
Прочие			210
Итого:			1600

6.2.3.6 Формирование бюджета затрат проекта

Бюджет проекта формируется из заработной платы, материальных затрат, амортизационных отчислений и отчислений во внебюджетные фонды. Итог расчет бюджета реализации исследования приведен в Таблице 6.7. Для большей наглядности структура бюджета проекта типизации геолого-

геофизического разреза с применением нейронных сетей представлена на Рисунке 6.2.

Таблица 6.7 – Бюджет затрат НТИ

№	Наименование статьи	Сумма, руб.	В % к итогу
1	Материальные затраты	1600,00	0,7
2	Амортизационные отчисления	7556,62	3,4
3	Затраты по заработной плате исполнителей темы	131358.30	58,6
4	Отчисления во внебюджетные фонды	41771.94	37,3
	Бюджет проекта	224058.8	100,00

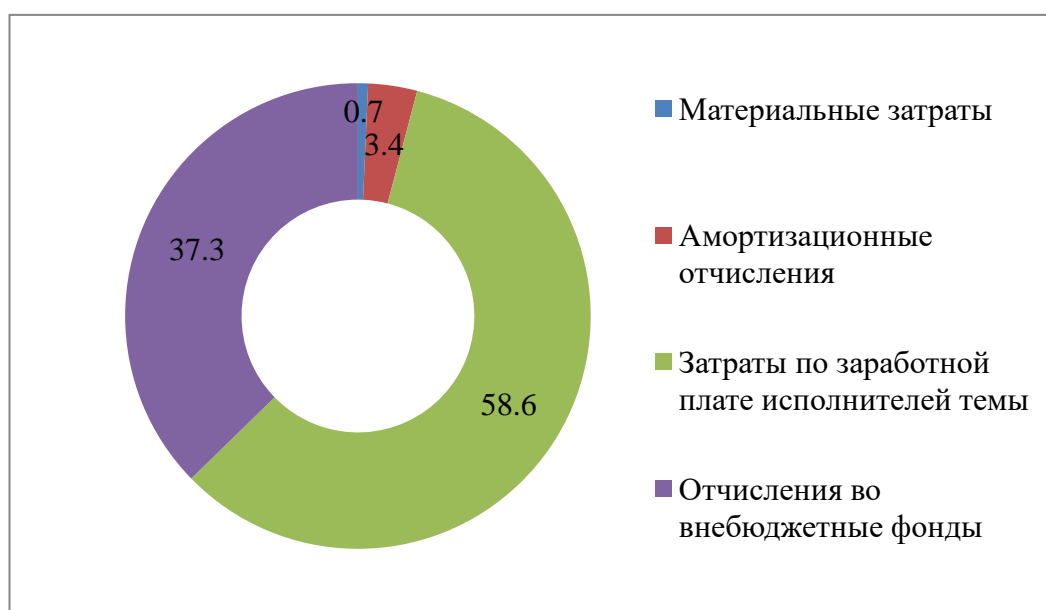


Рисунок 6.2 – Структура бюджета проекта в процентах

6.2.4 Реестр рисков проекта

Риск проекта – совокупность событий, явлений и их связей, которые могут привести к невозможности реализовать проект. Все риски можно подразделить на две группы: известные и неизвестные. Так риски, отнесенные к группе известных, – это прогнозируемые риски, влияние которых на проект должно быть учтено еще до начала его реализации. К такому типу рисков можно отнести влияние цены на нефть на спрос на программные продукты по типизации геолого-геофизических разрезов. Такие риски закладываются в проект и при наступлении того или иного сценария проект адаптируется по заранее высчитанной стратегии.

Однако существуют риски, вероятность возникновения которых оценить невозможно, такие риски относят к неизвестным рискам. Для противостояния таким рискам, обычно создаются резервы ресурсов, которые могут быть вовлечены в проект в наименьшее время и смогут нивелировать сложившуюся угрозу для проекта. В Таблице 6.8 указаны основные возможные риски, также экспертным образом оценена вероятность их возникновения и дана оценка их влияния по шкале от 1 до 5.

Таблица 6.8 – Выявленные риски проекта

№	Риск	Потенциальное воздействие	Вероятность риска	Влияние риска	Способы смягчения риска	Условия наступления
1	Снижение спроса на программное обеспечение для типизации геолого-геофизических разрезов	Снижение дохода от реализованных программных комплексов	4	4	Снижение себестоимости программы, выход на международные рынки	Снижение цены на нефть
2	Возникновение конкурентной программы	Снижение спроса на ПО	3	5	Позиционирование как ПО более лучшее. Доработка ПО.	Использование конкурентом более точных алгоритмов
3	Обесценивание национальной валюты	Потеря реальной прибыли	2	3	Производить операции в устойчивых валютах.	Политическое давление на РФ/падение цены на нефть/мировой кризис
4	Банкротство банка, где был открыт счет	Потеря денежных средств, возникновение задолженности	3	2	Открытие счетов в нескольких банках, с наилучшей репутацией.	Большое количество невозвратных кредитов

6.3 Определение эффективности исследования

Определение эффективности реализации проекта состоит в калькуляции интегрального показателя эффективности, который рассчитывается исходя из двух средневзвешенных критериев эффективности: финансовой и ресурсной.

Интегральный показатель финансовый эффективности проекта оценивается из сравнения нескольких методов реализаций поставленной задачи. Так в самом начале раздела анализировались два метода решения задачи типизации геолого-геофизических разрезов. Так при применении ручного метода специалист геолог-геофизик должен вручную обработать большой массив геофизических данных, для достижения такой же высокой точности реализации корреляции разрезов нескольких скважин. Так на анализ данных по одной новой скважине приходится порядка 4 часов рабочего времени специалиста, что составляет половину его рабочего времени. Учитывая амортизацию оборудования, зарплату специалиста, отчисления в фонды и материальные затраты стоимость обработки одной скважины обойдется в сумму порядка 1350 рублей, 750 из которых зарплата специалиста за половину рабочего дня. При реализации данного проекта аналогичная работа займет у специалиста порядка 5 минут с учетом запуска программы и выбора необходимого файла загрузки. Затраты на обработку при этом составят порядка 18 рублей, с учетом оплаты специалисту 10 рублей за 5 минут рабочего времени.

Интегральный финансовый показатель может быть определен как:

$$I_{\text{фин}} = \frac{\Phi_{pi}}{\Phi_{\text{max}}}, \quad (8)$$

где $I_{\text{фин}}$ – интегральный финансовый показатель; Φ_{pi} – стоимость исполнения исследуемого варианта; Φ_{max} – максимальная стоимость исполнения проекта.

Полученная величина $I_{\text{фин}}$ отражает во сколько раз данный вариант реализации проекта финансово эффективнее.

$I_{\text{фин}} = \frac{18}{1350} = 0,013$ – интегральный финансовый показатель реализации проекта типизации геолого-геофизических разрезов, показывает, что затраты при применении решений проекта снизятся на 98.7 %.

Интегральный показатель ресурсной эффективности можно определить следующим образом:

$$I_{pi} = \sum_1^n a_i \cdot b_i \quad (9)$$

где I_{pi} – интегральный показатель ресурсной эффективности проекта для i -го варианта исполнения разработки; a_i – весовой коэффициент i -го исполнения разработки; b_i – экспертная оценка i -го варианта исполнения разработки, по выбранной шкале оценивания; n – число параметров сравнения.

Произведенный расчет интегрального показателя ресурсной эффективности приведен в Таблице 6.9.

Таблица 6.9 – Сравнительная характеристика вариантов исполнения проекта по ресурсной эффективности

Критерий	Весовой коэффициент параметра	Ручной	Использование нейронных сетей
Скорость выполнения	0,25	1	5
Точность выполнения	0,25	5	5
Верность корреляции	0,25	5	4
Возможность внесения изменений	0,05	2	5
Наглядность полученного результата	0,05	3	5
Удобство использования результатов	0,05	1	5
Возможность добавления новых данных	0,1	2	5
Интегральный показатель:		3,25	4,75

Интегральный показатель эффективности вариантов исполнения разработок $I_{\text{исп}i}$ определяется на основании интегрального показателя ресурсоэффективности и интегрального финансового показателя по формуле:

$$I_{\text{исп}i} = \frac{I_{pi}}{I_{\text{фин}}}. \quad (10)$$

Сравнительная эффективность проекта оценивается как $\mathcal{E}_{\text{ср}} = \frac{I_{\text{исп}1}}{I_{\text{исп}2}}$ и приведена в Таблице 6.10.

Таблица 6.10 – Сравнительная эффективность проекта

№ п/п	Применяемые показатели	Ручной	Использование нейронных сетей
1	Интегральный финансовый показатель	1	0.013
2	Интегральный показатель ресурсной эффективности	3,25	4,75
3	Интегральный показатель эффективности	3.25	365.4
4	Сравнительная эффективность вариантов исполнения	0.01	112,43

Согласно расчётным данным, приведенным в Таблице 6.10, можно сделать вывод, что реализация геолого-геофизическая типизация разреза с применением нейронной сети, очевидно, дает огромное преимущество с точки зрения ресурсной и финансовой эффективности проекта.

7 Социальная ответственность

Целью данной магистерской диссертации является применение искусственной нейронной сети для типизации геолого-геофизического разреза. Данная работа выполнялась с применением методов машинного обучения. Для ее выполнения рабочее место должно состоять из следующих компонентов: рабочий стол, стул, персональный компьютер, с расположенным монитором на столе.

Данная работа выполнялась в учебной аудитории 19-го корпуса ТПУ №222. Учебное помещение с 9-ю рабочими местами имеет длину в 6.5 метров и ширину в 5 метров. Принципиальная схема рабочего помещения представлена на Рисунке 7.1.



Рисунок 7.1 – Принципиальная схема рабочего помещения

Персональные компьютеры, расположенные в данном помещении, располагаются на столах, принципиальная схема рабочего места будет рассмотрена далее в разделе. Компьютеры оборудованы процессором AMD Phenom™ II X2 555 Processor 3.2 ГГц, операционной памятью 5 Гб, видеоадаптером ATI Radeon HD 4250. Монитор компьютера BENQ GW2470 имеющий диагональ 24 дюйма, работающий с настройками 70% яркости и 50% параметра контрастности цветов. Установленный монитор работает с частотой 60 Гц при разрешении 1920x1080 FHD.

В данном разделе магистерской диссертации производится анализ рабочего места и рабочей зоны на соответствие требования законодательства РФ в области охраны и организации труда. Для этого необходимо проанализировать опасные и вредные факторы, возникающие при работе, а также разработать необходимый комплекс мер и мероприятий, для снижения негативного влияния, оказываемого на человека, общество и природу.

7.1 Производственная безопасность

В ходе выполнения данной магистерской диссертации использовался персональный компьютер. При работе с компьютером возникают вредные и опасные производственные факторы, которые приведены в Таблице 7.1.

Таблица 7.1 – Факторы, возникающие при работе с ПК

Источник фактора, наименование видов работ	Факторы		Нормативные документы
	Вредные	Опасные	
<ul style="list-style-type: none"> - Анализ входной информации в виде каротажных диаграмм - Создание программы позволяющей производить классификацию -Произведение анализа результатов -Создание отчетности 	<ul style="list-style-type: none"> 1) недостаточное освещение; 2) нарушение параметров микроклимата; 3) ионизирующие излучения; 4) умственное перенапряжение и монотонность работы. 	<ul style="list-style-type: none"> 1) поражение электрическим током; 2) статическое электричество; 3) короткое замыкание. 	<ul style="list-style-type: none"> 1) СанПиН 2.2.4.548-96; 2) СанПиН 2.2.4.5482013–96 3) СанПиН 2.2.2/2.4.1340-03; 4) СП 52.13330.2011; 5) ГОСТ Р 12.1.019-2009 ССБТ; 6) СНиП 2.01.02-85. 7) Противопожарные нормы.

Часто при постоянной работе с компьютером у работника возникает ряд профессиональных заболеваний, вызванными вредными факторами производства. Так наиболее характерными профессиональными заболеваниями, возникающих у работников, чей труд связан с каждодневным обращением с компьютером, являются: миопия глаз (близорукость), астигматизм, синдром сухого глаза, остеохондроз, гиподинамия, туннельный синдром, аллергия и нервные расстройства [21]. Влияние приведенных вредных факторов подрывает здоровье работника, влияя не только на физические показатели человека, но также и на его психосоматическое состояние. В совокупности данные вредные производственные факторы могут привести к производственной травме, что в целом говорит о имеющейся угрозе для жизни и здоровья человека.

7.1.1 Вредные факторы

7.1.1.1 Освещение

Типизация геолого-геофизического разреза подразумевает постоянную работу с каротажными диаграммами. Визуализация каротажных кривых необходима для произведения классификации и выявления реперных горизонтов в разрезе. Работа связана с постоянным напряжением зрительных органов. Как уже было сказано при анализе совокупности имеющихся вредных факторов, возникающих при работе, типичными заболеваниями являются болезни органов зрения. Для сдерживания и снижения влияния данного фактора необходимым условием является достаточное освещение рабочего места. Правильное освещение не только снизит напряжение глаз при работе, но и как следствие этого снижения увеличит производительность.

Помещение аудитории №222 19-го корпуса ТПУ спроектировано без окон. Таким образом доступа естественного освещения в аудитории нет. Согласно СНиП 23-05-95 «Естественное и искусственное освещение» [22] в таких помещениях допускается работа не выше IV категории точности (средняя точность). Поэтому в рабочем помещении используется лишь искусственное

освещение. Согласно СанПиН 2.2.2/2.4.1340—03 искусственное освещение в помещениях для эксплуатации ПЭВМ должно осуществляться системой общего равномерного освещения [23]. На Рисунке 7.2 приведена схема освещения рабочего помещения, в котором установлены светильники типа ARS/R 418 размером 500x500x40 мм, каждый укомплектован четырьмя люминесцентные лампы типа Philips TL-D, каждая мощностью 18 Вт.

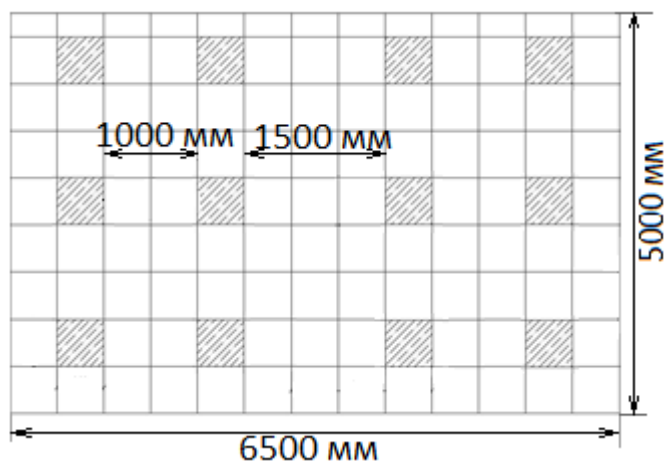


Рисунок 7.2 – Схема освещения помещения

Помещение имеет размерность: длина 6,5 м, ширина – 5 м, высота – 2,91 м, высота стола – 0,76 м, суммарное количество ламп составляет 48 штук. Согласно [22] норма освещённости рабочих поверхностей в помещениях для работы с дисплеями и видеотерминалами принимается $E_n = 400$ лк.

Фактическая освещенность определяется по формуле:

$$E_{\phi} = \frac{N \cdot n \cdot \Phi_{\text{ст}} \cdot \eta}{S \cdot K_z \cdot Z}, \quad (11)$$

где N – число светильников; n – число ламп в светильнике; $\Phi_{\text{ст}}$ – световой поток люминесцентных ламп, лм; η – коэффициент использования светового потока, %; S – площадь освещаемого помещения, м^2 ; K_z – коэффициент запаса, учитывающий загрязнение светильника; Z – коэффициент неравномерности освещения.

Для определения коэффициента использования светового потока необходимо найти индекс помещения. Индекс помещения определяется по формуле:

$$i = \frac{S}{h \cdot (A+B)}, \quad (12)$$

где h – высота от ламп до рабочей поверхности, A – ширина, а B – длина. Рассчитаем индекс:

$$i = \frac{6,5 \cdot 5}{(2,91 - 0,76) \cdot (6,5 + 5)} \approx 1,31. \quad (13)$$

При $i = 1,31$, коэффициенте отражения светлых стен $\rho_c = 50\%$ и светлого потолка $\rho_n = 70\%$, коэффициент использования светового потока, согласно [22], равен $\eta = 0,48$. Коэффициент неравномерности освещения Z для люминесцентных ламп примем $Z = 1,1$. Коэффициент запаса в случае офисных помещений составляет $K_z = 1,2$. Световой поток используемых люминесцентных ламп составляет 1200 лм. Рассчитаем фактическую освещенность:

$$E_\phi = \frac{12 \cdot 4 \cdot 1200 \cdot 0,48}{6,5 \cdot 5 \cdot 1,2 \cdot 1,1} \approx 644,48 \text{ лк}. \quad (14)$$

Рассчитаем отклонение от нормы ΔE :

$$\Delta E = \frac{750 - 644,48}{750} \cdot 100 \approx 14,1\%. \quad (15)$$

Так фактическая освещенность на 14,1 % ниже, чем нормативное значение для помещений, предназначенных для работ средней точности, при условии постоянной работы в нем. Учитывая, что при периодической работе с постоянным пребыванием норма составляет 500 лк, можно указать на соответствие освещенности учебной аудитории требуемому уровню. При необходимости обеспечения постоянной работы в таком помещении необходимо поменять люминесцентные лампы на лампы, генерирующие больший световой поток, либо установить дополнительные источники освещения.

Данное рабочее место, используемое на время написания магистерской диссертации, соответствует нормативным требованиям освещенности.

7.1.1.2 Микроклимат

Одной из важнейших характеристик пригодности рабочей зоны для выполнения работ определенной категории является микроклимат

Согласно ГОСТ 12.1.005-88 ССБТ под микроклиматом производственных помещений принято понимать метеорологические условия внутренней среды помещений, которые определяются действующими на организм человека сочетаниями температуры, влажности, скорости движения воздуха и теплового излучения [24].

Данные характеристики микроклимата должны обеспечивать сохранение теплового баланса между человеком и окружающей его средой, а также поддерживать оптимальное или допустимое тепловое состояния организма. СанПиН 2.2.4.5482013-96 содержит основные нормативные требования, применяемые к показателям микроклимата и их сочетаниям. Нормы указанного СанПиНа разработаны на основе исследований теплообмена и теплопереноса, а также теплового состояния человека, находящегося в микроклиматических камерах. Данные камеры моделируют условия производственных помещений. Также СанПиН 2.2.4.5482013-96 содержит результаты клинических и санитарно-эпидемиологических исследований [25].

В Таблице 7.2 приведены оптимальные и допустимые показатели микроклимата на рабочих местах в помещениях.

Таблица 7.2 – Допустимые и оптимальные показатели микроклимата

Период года	Температура, °С		Относительная влажность, %		Скорость движения воздуха, м/с	
	опт.	доп.	опт.	доп.	опт.	доп.
Холодный	21 – 23	22 – 24	30 – 50	65	не более 0,2	не более 0,2
Теплый	22 – 24	23 – 27	30 – 50	60	не более 0,2	не более 0,2

Так как учебная аудитория № 222 учебного корпуса №19 ТПУ не имеет окон, поступление воздуха в помещение происходит только по средствам

вентиляционной системы. Необходимые меры по улучшению воздушной среды в рабочем помещении:

- правильная организация вентиляции и кондиционирования воздуха;
- отопление помещений.

В учебной аудитории с объемом 94.6 м³ подается объем наружного воздуха в размере 45 м³/час на человека. В условиях отсутствия естественной вентиляции, подача воздуха и контроль температуры входящего воздуха осуществляются с помощью центральной системы подачи и кондиционирования воздуха. В зимний период в рабочем помещении предусмотрена система водяного центрального отопления со встроенными нагревательными элементами и терморегуляторами. Таким образом, данное рабочее место с точки зрения микроклиматических условий полностью соответствует нормативным требованиям законодательства РФ в сфере охраны и организации труда.

7.1.1.3 Ионизирующие излучения

Под ионизирующими излучениями понимают такие виды лучистой энергии, которые, попадая в определенные среды или проникая через них, производят в них ионизацию, т.е. превращают атомы и молекулы в ионы.

В настоящее время используемые жидкокристаллические мониторы производят настолько низкое ионизирующее излучение, что оно является ничтожно малым и практически не превышает естественный уровень излучения. Однако дешевые низкокачественные и бракованные мониторы могут быть источниками излучения, превышающего допустимые нормы, в данном случае необходимо использовать качественные исправные мониторы, желательно, известных торговых марок.

Согласно СанПиН 2.6.1.2523-09 [26] для отнесения условий труда при обращении с техногенными и природными источниками излучения на производстве к допустимым, максимальная потенциальная эффективная доза не должна превышать 5 мЗв/год, а максимальная эквивалентная доза в хрусталике

глаза, коже, кистях и стопах не должна превышать 37,5, 125 и 125 мЗв/год, соответственно.

Также СанПиН 2971-84 [27] устанавливает правила защиты от воздействия электрических полей и устанавливает предельно допустимый уровень напряжённости электрического поля 0.5 кВ/м внутри зданий. Согласно ГОСТ 27016-86 [28] плотность потока ультрафиолетового излучения дисплея не должна превышать показания в 10 Вт/м², а доза рентгеновского излучения при расстоянии в 5 см от экрана монитора - не больше 0,03 мкР/с.

Для снижения негативного влияния, оказываемого ЖК-монитором необходимо применить следующие меры:

- выключать монитор в случае ненадобности (перерыв или конец рабочего дня);
- применять специальные защитные экраны и/или покрытия;
- строго соблюдаются общие правила работы за компьютером;
- рабочее место должно быть убрано, а помещение проветриваемо;
- применять устройства увлажняющие и/или ионизирующие воздух.

Анализируя рабочее помещение, а именно учебную аудиторию № 222 учебного корпуса №19, следует сделать вывод о соответствии данного помещения требуемым нормам в части защиты от ионизирующего излучения.

7.1.1.4 Психосоматические аспекты работы

Выполнение данной магистерской диссертации связано с использованием компьютера. Разрабатываемый метод типизации геолого-геофизического разреза также неотрывно связан с вредным воздействием электронной техники. Специалистам геологам и геофизикам сейчас необходимо затрачивать гораздо большее время работы за компьютером при решении задачи типизации без применения алгоритмов машинного обучения. Негативное влияние оказывается не только на физическую активность человека. Очень часто монотонный труд и длительное напряжение глаз вызывает головные боли. Частое напряжение тела при работе за компьютером,

неправильно подобранная офисная мебель вызывает зажатость артерий, сосудов и как следствие уменьшение притока свежей крови в органы. Часто от перенапряжения возникают головные боли, возникают стрессовые состояния, может ухудшаться сон. Монотонность работы без применения методов машинного обучения приводит к снижению концентрации внимания. Данные негативные факторы уменьшаются при использовании для решения задачи типизации геолого-геофизических разрезов методов машинного обучения. Методы машинного обучения позволяют ускорить процесс работы, при уменьшении участия человека. Нет необходимости в монотонном сравнении кривых каротажей человеком. Однако негативное воздействие электронной техники остаются, сохраняется и возникающая на этом фоне усталость, а также стресс работника.

Чтобы снизить негативное воздействие необходимо соблюдать составленный в соответствии с СанПиН 2.2.2/2.4.1340-03 [23] распорядок рабочего дня. Распорядок рабочего дня при работе с компьютером включает промежутки отдыха в рабочее время, что позволяет уменьшить нагрузку органов зрения, а также снижению стрессовых ситуаций. Следует отметить, что согласно ТОО Р-45-084-01 [29] по характеру работы, такой вид деятельности, как типизация геолого-геофизических разрезов, относиться к группе «Б». Данная группа накладывает ограничение по суммарному числу считываемых или вводимых знаков за рабочую смену (не более 40000 знаков за смену).

Так при 8-ми часовой рабочей смене сотрудник обязан сделать перерывов на 70 минут в течении дня, а при 12-ти часовой смене на 120 минут. При этом перерывы должны быть через 1,5 - 2,0 часа от начала рабочей смены и через 1,5 - 2,0 часа после обеденного перерыва продолжительностью 20 минут каждый или продолжительностью 15 минут через каждый час работы. Продолжительность работы за компьютером без перерывов не должна превышать двух часов.

7.1.2 Опасные факторы

7.1.2.1 Поражение электрическим током

При работе с компьютером возникает опасность поражения электрическим током пользователя. Необходимо, чтобы каждый работник соблюдал правила требования безопасности при работе на ПК, установленные [29]. Данные правила устанавливают следующие требования:

- держать рабочее место в чистоте и порядке;
- перед началом работы убедиться в подключении оборудования к электросети;
- проверить правильность подключения и отсутствия оголенных соединений, проводов;
- убедиться в наличии заземления блока питания;
- использовать антистатические салфетки для обработки монитора компьютера.

Также работнику запрещается:

- прикасаться к задней панели блока питания, при включенном питании;
- переключать разъемы интерфейсного оборудования при включенном питании;
- допускать попадания влаги на поверхность монитора, блока питания, клавиатуры и принтеров;
- производить самостоятельный ремонт и вскрытие техники;
- работать за компьютером при снятых защитных кожухах системного блока;
- отключать оборудование от сети и выдергивать вилку держась за шнур.

Помещение, где расположено рабочее место пользователя ПЭВМ, относится к помещениям без повышенной опасности ввиду отсутствия следующих факторов:

- сырость (относительная влажность воздуха превышает 75%);
- токопроводящая пыль;
- высокая температура (превышает 35 °С);

- токопроводящие полы;
- возможность одновременного прикосновения человека к имеющим соединение с землей металлоконструкциям зданий, технологическим аппаратам, механизмам и металлическим корпусам электрооборудования.

Электрический ток имеет три вида воздействия на организм человека: термический, электролитический, биологический. Последствиями воздействия электрического тока на организм являются: электрический удар, электрическая травма, прекращение работы сердца, остановка дыхательных процессов. Таким образом, воздействие электрическим током может привести не только к различного рода травмам, но и к летальному исходу. Именно поэтому электрический ток является одним из опаснейших факторов воздействия.

Особенно необходимо отметить обязательность проведения организационно-технических инструктажей безопасности для сотрудников. Следует заключить, что рабочее место соответствует требованиям законодательства РФ в области охраны и организации труда в части защиты от поражения электрическим током.

7.1.2.2 Статическое электричество

Одной из часто возникающих проблем при работе за ПК связанных с электрическим током, является статическое электричество. Статическое электричество – это физическое явление, связанное с возникновением свободного электрического заряда на поверхности или внутри диэлектрических материалов. В процессе работы компьютер, как электронный прибор генерирует электростатическое поле, при этом в зону действия электростатического поля попадают периферийные устройства и монитор, с которыми взаимодействует пользователь. Наэлектризованные полем частички пыли способны оседать на поверхности кожи и в тканях дыхательных путей человека, что способно вызвать изменение ионного состава тканей. Длительное воздействие электростатических полей вызывает повышенную утомляемость и раздражительность человека.

Именно для защиты от электростатического поля необходимо наличие заземления блока питания, также необходимо протирать монитор специальными антистатическими салфетками, устанавливать защитные экраны. Не менее важно, держать рабочее место в чистоте, для избегания наэлектризованности частичек пыли.

Таким образом, рабочее место соответствует требованиям законодательства РФ в области охраны и организации труда в части защиты от влияния статического электричества.

7.1.2.3 Короткое замыкание

Наиболее опасным при работе с ПЭВМ является возникновение короткого замыкания. Короткое замыкание – представляет собой соединение двух точек электрической цепи с различными значениями потенциала, не предусмотренное конструкцией устройства и нарушающее его нормальную работу. При возникновении короткого замыкания сила тока возрастает мгновенно и достигает высоких значений, что приводит к быстрому росту тепловыделения, расплавлению частей электронного устройства и электрических проводов, что приводит к возникновению пожара.

Наиболее частой причиной короткого замыкания является нарушение изоляции электроприборов. В свою очередь это может быть вызвано: перенапряжениями сети, прямыми ударами молнии, старением оборудования и изоляции, механическими повреждениями, неправильным уходом или эксплуатацией устройства, несоблюдением техники безопасности при работе и др.

Для предотвращения возникновения короткого замыкания и снижения вероятности пожара, а также получения работником электротравмы, электрооборудование должно быть оборудовано защитным заземлением (занулением), защитными кожухами, предупредительными указателями в соответствии с техническими требованиями по эксплуатации. Устройства, подключенные в сеть высокого напряжения, должны быть подключены не

напрямую, а только через сетевые фильтры или розетки с заземляющей шиной. Электронно-вычислительная техника должна быть подключена с использованием сетевых фильтров, которые нивелируют импульсные скачки тока в сети, перегрузки и предотвращают короткое замыкание.

Согласно законодательству РФ в области охраны и организации труда, в части защиты от короткого замыкания рабочее место соответствует предъявляемым требованиям.

7.2 Экологическая безопасность

Типизация геолого-геофизических разрезов как таковая не имеет негативного воздействия на окружающую среду, так как является интеллектуальной деятельностью. Однако средства, используемые для ее достижения, могут быть рассмотрены, как наносящие такой вред.

В качестве промышленных отходов выступают: бумага и пластик. К бумажным отходам могут быть отнесены записки, пометки, документация, бразовавшаяся в ходе выполнения проекта. Пластиковые отходы могут образовываться при поломке техники. НИ ТПУ осуществляет сбор и передачу бумажных отходов сторонним компаниям для переработки и повторного использования. Пластиковый мусор требует более сложных манипуляций, часть отходов может содержать вредные вещества, поэтому их необходимо утилизировать, только в сертифицированных центрах по переработке вторичного сырья.

Люминесцентные лампы осветительных приборов должны быть подвергнуты особому процессу транспортирования и утилизации, согласно ГОСТ 6825-91 [30]. Утилизировать такого рода отходы имеет право компания, обладающая лицензией на осуществление такой деятельности.

7.3 Безопасность в чрезвычайных ситуациях

Согласно определению чрезвычайная ситуация (ЧС) – обстановка сложившаяся на определенной территории, возникшая в результате аварии,

природного явления, катастрофы, стихийного или иного бедствия, которые могут повлечь или повлекли за собой человеческие жертвы, ущерб здоровью людей или окружающей природной среде, значительные материальные потери и нарушение условий жизнедеятельности людей [31]. По своему типу ЧС могут быть классифицированы как: техногенные, природные, экологические, биологические и антропогенного характера.

Наиболее вероятно при работе с ПЭВМ возникновение ЧС техногенного характера, такие как пожары и взрывы. Пожарная безопасность, как единый комплекс мер по обеспечению безопасности и предотвращению ЧС, является особенно важным аспектом деятельности предприятий.

Согласно [32], рабочее помещение относится к категории В, согласно наличию в нем электронной техники. Источниками повышенной пожароопасности являются:

- электрооборудование без защиты от короткого замыкания;
- неисправная аппаратура;
- нарушение правил эксплуатации оборудования.

Профилактические меры необходимые для снижения вероятности возникновения пожара, а также снижения угрозы жизни и здоровью работников при возникновении ЧС:

- первичный и повторный инструктаж персонала и разработка планов эвакуации;
- проверка наличия схем эвакуации и первичных средств пожаротушения;
- периодические проверки исправности электротехнического оборудования;
- периодические проверки эвакуационных выходов;
- профилактическое отключение оборудования;
- установка средств пожарной сигнализации;
- использование негорючих материалов;
- установка устройств ограничения распространения огня в помещениях;
- установка стационарных систем тушения;

– ревизия и контроль состояния систем тушения и защиты от огня.

При возникновении пожара необходимо, в первую очередь, незамедлительно сообщить о пожаре в единую службу спасения, также предпринять меры по эвакуации людей в соответствии с планом эвакуации. Только в случае, если отсутствуют прямые угрозы здоровью и жизни людей произвести попытку самостоятельного тушения возникшего возгорания с применением средств первичного пожаротушения. Необходимо строго понимать какое средство можно применить для тушения конкретного типа возгорания. Так тушение электрических приборов приведет к удару электрическим током и усугубит сложившуюся ЧС. При потере контроля ситуацией, необходимо срочно эвакуироваться согласно плана эвакуации и ждать приезда пожарной службы. Важно встретить пожарную бригаду, сообщить ей имеющуюся информацию о источнике возгорания, о местонахождении людей в здании, о наличии взрывоопасных материалах, а также о конструкции здания (проходы, тупики и т.д.).

Рассматриваемое помещение в данном случае оборудовано пороговой пожарной сигнализацией, автоматическими противопожарными системами, пожарными гидрантами и огнетушителями. Знаки пожарной безопасности нанесены и четко различимы, схемы эвакуации присутствуют. В целом рабочее место соответствует нормам пожарной безопасности.

7.4 Правовые и организационные вопросы обеспечения безопасности

Трудовой кодекс РФ является основополагающим документом в регулировании отношений между работодателем и работником, в части оплаты труда, трудового распорядка, а также особенности регулирования труда женщин, детей, людей с ограниченными способностями.

Для обеспечения комфортных условий труда и безопасности сотрудников в соответствии с нормами законодательства РФ, необходимо

обеспечить рациональную организацию трудовой деятельности персонала, что включает в себя:

- длительность рабочей смены не должна превышать 8 часов;
- запрещена непрерывная работа за компьютером более 2 часов;
- необходимо наличие, как минимум, 2 регламентируемых перерывов (20 минут после 1,5 – 2 часов работы, либо по 15 минут после каждого рабочего часа);
- длительность работы с ПЭВМ не должна превышать 6 часов за рабочую смену;
- длительность обеденного перерыва должна составлять не менее 40 минут.

ГОСТ Р ИСО 9241-1-2007 [33] регламентирует вопросы эргономичности рабочей зоны и рабочих мест. В свою очередь в ГОСТ Р ИСО 9241-2-2009 [34] устанавливаются требования к конфигурации рабочих станций, которые имеют цель обеспечить комфортные условия для работы сотрудника, а также безопасность выполняемой работы.

СанПиН 2.2.2/2.4.1340-03 [23] устанавливает норму полезной площади помещений для сотрудников различного рода служб. Так на работников, чья профессиональная деятельность связана с использованием МЭВМ должно приходиться не менее 4 м² полезной площади рабочего помещения.

Вопрос организации рабочего места сотрудника, работающего с ПК, является основополагающим в организации комфорта и безопасности работы. Так при оборудовании рабочего места сотрудника монитор должен располагаться на расстоянии в 60–70 см от глаз работника. Особое внимание должно быть уделено рабочему креслу, так как сотрудник проводит в нем большую часть рабочей смены оно должно обеспечивать правильную осанку и обеспечивать комфортность при работе. Конструкция рабочего кресла, входящего в состав рабочего места должна создавать поддержку правильной позиции спины при работе, а также позволять менять позу, для обеспечения снижения напряжения мышц спины и плечевого пояса. Рабочее место должно обеспечивать удобство и безопасность при работе, также оно должно быть

регулируемым для соответствия индивидуальным требованиям работника, ввиду его биометрических параметров (рост, вес, рабочая рука и др.). Пример рабочего места используемого при создании данной магистерской диссертации приведен на рисунке 7.3.

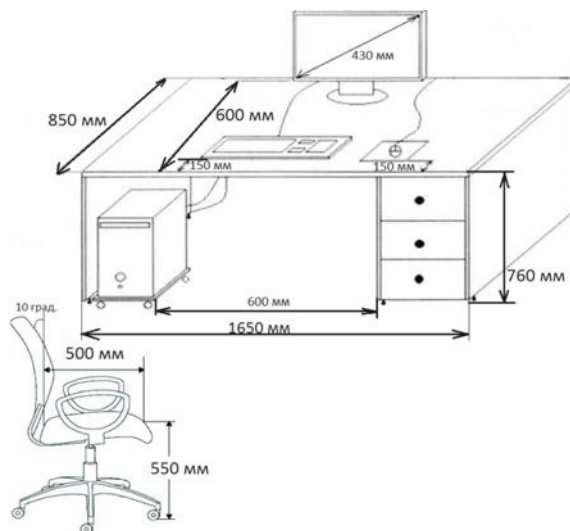


Рисунок 7.3 – Основные характеристики используемого рабочего места

Данное рабочее место соответствует приведенным в разделе нормам и законодательным актам и в целом удовлетворяет предъявляемым требованиям по обеспечению комфорта и безопасности работников.

Заключение

В заключении следует отметить, что решение задачи типизации геолого-геофизических разрезов с применением нейронных сетей является нетривиальной задачей. Решения, полученные при реализации данного проекта, имеют высокую точность и ускоряют процесс обработки данных геофизических исследований.

Актуальность проведенного исследования заключается в том, что методы машинного обучения позволяют производить интерпретацию каротажных данных автоматически с приемлемой точностью.

Для решения задачи классификации отложений горных пород были использованы 5 различных классификаторов, два из которых дали наиболее точные результаты. А именно классификатор на основе искусственной нейронной сети, а также классификатор на основе каскада сетей. Наилучшим был признан классификатор, на основе каскада сетей дающий общую точность в 88.7% при наиболее высокой скорости обучения классификатора.

Задача выделения реперных горизонтов была решена с применением классификатора на основе скрытых Марковских моделей. Данный тип классификатора наилучшим образом подходит для обработки многомерных последовательностей различных мощностей, которыми являются реперные горизонты. Следует отметить, что его преимущество заключается в хорошей точности, даже при обучении на малой обучающей выборке.

В целом точность полученных результатов при решении поставленных задач достаточно высока, однако планируется дальнейшая доработка разработанного программного комплекса.

Список публикаций студента

- 1 Канаев И.С. Методы интенсификации газоотдачи угольных пластов и их применение в России / Проблемы геологии и освоения недр: труды XX Международного симпозиума имени академика М.А. Усова студентов и молодых ученых, посвященного 120-летию со дня основания Томского политехнического университета, Томск, 4-8 Апреля 2016. - Томск: Изд-во ТПУ, 2016. - Т. 2 - С. 350-352.
- 2 Канаев И.С. Месторождения газа в акватории Баренцева моря. Штокмановское месторождение / Проблемы геологии и освоения недр: труды XVII Международного симпозиума имени академика М.А. Усова студентов и молодых учёных, посвященного 150-летию со дня рождения академика В.А. Обручева и 130-летию академика М. А. Усова, основателей Сибирской горно-геологической школы: в 2 т., Томск, 1-6 Апреля 2013. - Томск: Изд-во ТПУ, 2013. - Т. 1 - С. 734-736.
- 3 Kanaev, I. Technogenic influence on the environment from ancient times to the present / Творчество юных - шаг в успешное будущее: материалы VI Межрегиональной научной студенческой конференции с элементами научной школы имени профессора М.К. Коровина, Томск, 25-28 Ноября 2013. - Томск: Изд-во ТПУ, 2014. - С. 205-207.
- 4 Канаев, И.С. Особенности разработки нефтяных месторождений в акватории Баренцева моря / Творчество юных - шаг в успешное будущее: материалы V Межрегиональной научной студенческой конференции с элементами научной школы имени профессора М.К. Коровина, Томск, 20-24 Ноября 2012.

Список используемых источников

- 1 Tschannen, V., Delescluse, M., Rodriguez, M., and Keuper, J. "Facies classification from well logs using an inception convolutional network," 2017.
- 2 Bestagini, P., Lipari, V., Tubaro, S. and Milano, P. "A machine learning approach to facies classification using well logs"/ SEG International Exposition and 87th Annual Meeting / pp. 2137–2142, 2017.
- 3 Malvić, T., Velić, J., Horváth, J., and Cvetković, M. "Neural networks in petroleum geology as interpretation tools," Cent. Eur. Geol., vol. 53, no. 1, pp. 97–115, 2010.
- 4 Хаустов П.А., Тезисы диссертации на соискание ученой степени к.т.н. по теме: "Алгоритмы распознавания рукописных символов в условиях малой обучающей выборки", Томск, ТПУ, 2017.
- 5 Ma, Y.Z. "Lithofacies clustering using principal component analysis and neural network: Applications to wireline logs" Math. Geosci., vol. 43, no. 4, pp. 401–419, 2011.
- 6 Seth Willis Bassett, B. "Teaching a Computer Geology: Automated Lithostratigraphic Classification Using Machine Learning Algorithms," DMT Meet., vol. 17, 2017.
- 7 Gupta, I., S. Rai, C., C. H. Sondergeld, and D. Devegowda, "Rock Typing in Eagle Ford, Barnett, and Woodford Formations," Proc. 5th Unconv. Resour. Technol. Conf., no. October 2017, pp. 1–17, 2017.
- 8 Шайбаков Р.А., "Использование нейросетевого аппарата для идентификации границ геологических объектов," Молодой ученый, 168с., стр. 8-11, 2012.
- 9 Bougher, B. C. "Prediction of stratigraphic units from spectral co-occurrence coefficients of well logs," CSEG Annu. Conv., pp. 1–5, 2015.
- 10 Гафуров Д. О., Гафуров О. М., Конторович В. А. "Интерпретация данных геофизических исследований Талаканского нефтегазоконденсатного месторождения обучаемыми нейронными сетями, прогноз строения

- осинского горизонта,” Технологии сейсморазведки, 104с., стр. 85–92, 2014.
- 11 Калинина Л.М., Тезисы диссертации на соискание ученой степени к.г.-м.н. по теме: “Геологическое строение, условия формирования и нефтегазоносность келловей и Оксфорда Западной Сибири в области перекода морских отложений в континентальные,” Новосибирск, 2005.
- 12 Корсунь В.В., “Новые данные о геологическом строении и нефтегазоносности Урманско-Арчинской зоны в Томской области,” Нефтяное хозяйство–2006.№12, стр. 52–56.
- 13 Зиганшина Р. А., Сидельник А. В. “Взаимосвязь новейших сдвигов и геодинамики с нефтеносностью (на примере краснелинского свода и нюрольской впадины, западная сибирь),” Труды четвертой тектонической конференции,стр. 39–45, 2016.
- 14 Конторович В. А., “Сейсмогеологические критерии нефтегазоносности зоны контакта палеозойских и мезозойских отложений Западной Сибири (На примере Чузикско-Чижапской зоны нефтегазонакопления),” Геология и Геофизика–2007.№5, стр. 538–547.
- 15 Демин В.Ю., Наумов Ю.А., Абалтусов Н.В., “Борьба с поглощением бурового раствора при разбуривании карбонатного коллектора на Арчинском месторождении” Инженерная практика, 80с., pp. 70–75, 2015.
- 16 Воробьева С. А. Методы распознавания речи // Молодой ученый. – 2016. – №26. – С. 136-141.
- 17 Cesar de Souza. Sequence Classifiers in C# / [Электронный доступ] / CodeProject, 2013. – Режим доступа: <https://www.codeproject.com/articles/541428/sequence>, (дата обращения: 17.08.2018).
- 18 Ng A., Jordan M. On Discriminative vs. Generative Classifiers: A comparison of logistic regression and Naive Bayes // In Advances in Neural Information Processing Systems 14. 2002. pp. 841– 848.

- 19 Паламарь И.Н., Юлин С.С. Порождающая графическая вероятностная модель на основе главных многообразий // Тр. СПИИРАН. 2014. № 2. С. 227-247.
- 20 Романенко А.А. Выравнивание временных рядов: прогнозирование с использованием DTW / А.А. Романенко // Машинное обучение и анализ данных. – 2001. – № 1 [Электронный ресурс]. – Режим доступа: <http://jmla.org/papers/doc/2011/no1/Romanenko2011Dynamic.pdf> (дата обращения: 17.08.2018).
- 21 Савельев В.А. «Компьютерные» болезни и их профилактика / Компьютер пресс. – 2006. – №6. – С. 66-72.
- 22 СНиП 23-05-95. Естественное и искусственное освещение / М.: Минстрой России, 2016. – 103 с.
- 23 СанПиН 2.2.2/2.4.1340-03. Гигиенические требования к персональным электронно-вычислительным машинам и организации работы.– М.: Информационно-издательский центр Минздрава России, 2003.–18с.
- 24 ГОСТ 12.1.005-88 Система стандартов безопасности труда (ССБТ). Общие санитарно-гигиенические требования к воздуху рабочей зоны.– М.: Минздрав СССР, 1988. – 20 с.
- 25 СанПиН 2.2.4.548-96. Гигиенические требования к микроклимату производственных помещений. – М.: Информационно-издательский центр Минздрава России, 1997. – 11 с.
- 26 СанПиН 2.6.1.2523-09. Нормы радиационной безопасности НРБ-99/2009. – М.: Информационно-издательский центр Минздрава России, 2009.–70 с.
- 27 СанПиН 2971-84. Санитарные нормы и правила защиты населения от воздействия электрического поля, создаваемого воздушными линиями электропередачи переменного тока промышленной частоты.–М.: Минздрав СССР, 1984. – 8 с.
- 28 ГОСТ 27016-86. Дисплеи на электронно-лучевых трубках. Общие технические условия.–М.: Госстандарт СССР, 1986.–12 с.
- 29 ТОИ Р-45-084-01. Типовая инструкция по охране труда при работе на

- персональном компьютере. – М.: Стандартиформ, 2001. – 4 с.
- 30 ГОСТ 6825-91. Лампы люминесцентные трубчатые для общего освещения.–М.: Госстандарт СССР, 1991.–242 с.
- 31 ГОСТ Р 22.0.02-2016. Безопасность в чрезвычайных ситуациях. Термины и определения. – М.: Стандартиформ, 2016. – 10 с.
- 32 Свод Правил 12.13130.2009. Определение категорий помещений, зданий и наружных установок по взрывопожарной и пожарной опасности. – М.: МЧС России, 2009. – 35 с.
- 33 ГОСТ Р ИСО 9241-1-2007. Эргономические требования к проведению офисных работ с использованием видеодисплейных терминалов (VDTS). Общее введение. – М.: Стандартиформ, 2008. – 13 с.
- 34 ГОСТ Р ИСО 9241-2-2009. Эргономические требования к проведению офисных работ с использованием видеодисплейных терминалов (VDTS). Требования к расположению рабочей станции и осанке оператора. – М.: Стандартиформ, 2010. – 46 с.

Приложение А

Раздел 3

Data analysis

Студент

Группа	ФИО	Подпись	Дата
2ТМ61	Канаев Илья Сергеевич		27.06.18

Консультант – лингвист отделения иностранных языков ШБИП

Должность	ФИО	Ученая степень, звание	Подпись	Дата
доцент	Максютин К.Ю.	к.п.н		

Консультант отделения нефтегазового дела ИШПР

Должность	ФИО	Ученая степень, звание	Подпись	Дата
ведущий эксперт ЦППС НД ТПУ	Чернова О.С.	к.г.—м.н.		

The raw input logging data of each well is stored in .las file format. The data of the same well logging method in adjacent wells can have significantly different averages due to different measurement conditions and also include gaps and can be registered in different ranges. Let's list the main problems in the source of data:

- Different set of log data.
- Large number of gaps in log data.
- Small amount of data at shallow depth and large amount in the areas of interest, uneven distribution of classes.
- Different lengths of sections of the same log type for the same formations in different wells.
- Various conditions for conducting measurements (for example, differently calibrated instruments), so that the data at the same point have different numerical values.
- Noise and emissions in the data.

At the first stage of the data preparation procedure for each well, well logging data samples with discarded missed or incorrect (less than zero) values are separately compiled. Then, logging normalization is performed by calculating standardized estimates (z-estimates) to solve the problem of different measurement conditions. Standardized evaluation is a dimensionless statistical measure for comparing values of different dimensions or measurement scales.

Based on the assumption of the normality of the data distribution, the log data set $X = \{x_1, x_2, \dots, x_n\}$ with the mean value \bar{X} and the standard deviation S_x can be transformed into another set with zero mean value and a single standard deviation in such way that the converted values of z will be directly expressed in the deviations of the initial values from the mean, measured in units of standard deviation. Thus, each logging point is transformed according to the following formula: $x_i = \frac{x_i - \bar{X}}{S_x}$.

In figures A.1, A.2, A.3 obtained histograms of standardized log estimates from the data of four wells in the C field are illustrated.

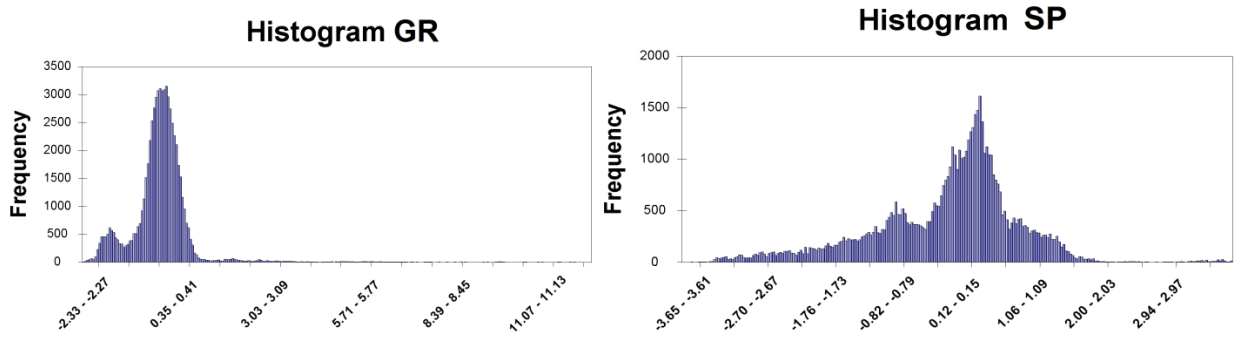


Figure A.1 – Histograms of standardized gamma ray log and self-potential log

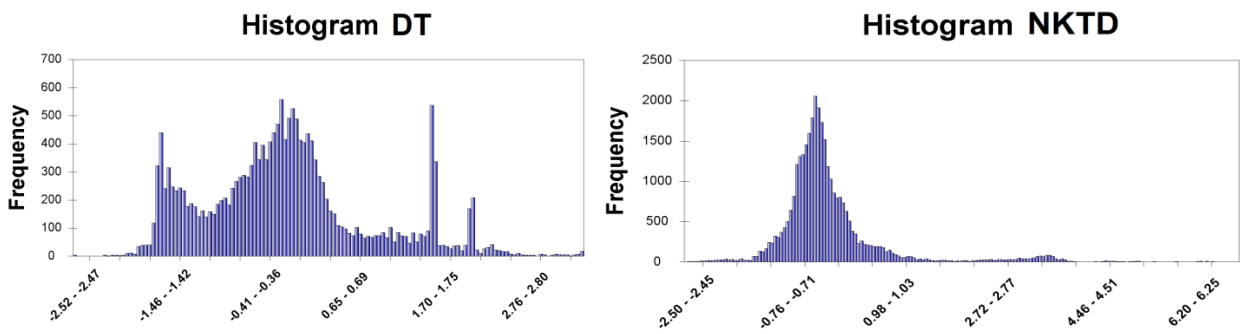


Figure A.2 – Histograms of standardized acoustic and neutron gamma logs

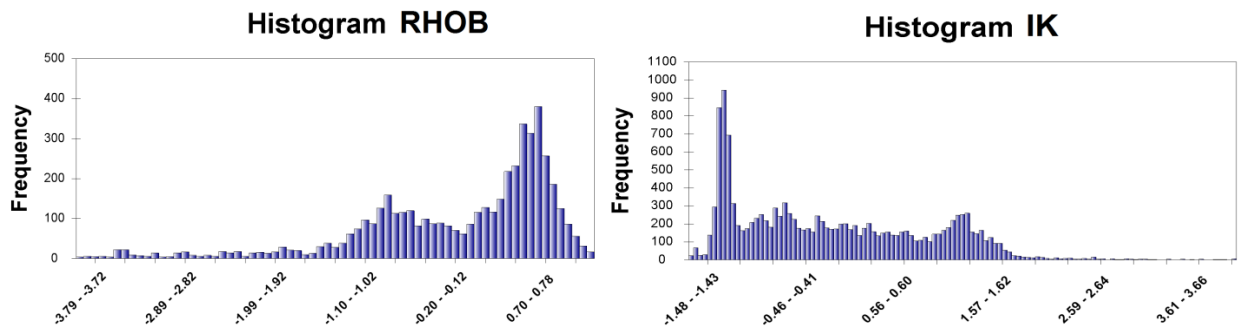


Figure A.3 – Histograms of standardized data of density and induction logs

Visualization of the initial data at the stage of data preprocessing procedure helps to better understand the initial regularities in the data: for example, the distribution of induction and density logs is lognormal. The transformation of input data in accordance with the distributions allows to improve the quality and reduce the learning speed of artificial neural networks. Figure A.4 illustrates the histogram of classes` (formation) data distribution.

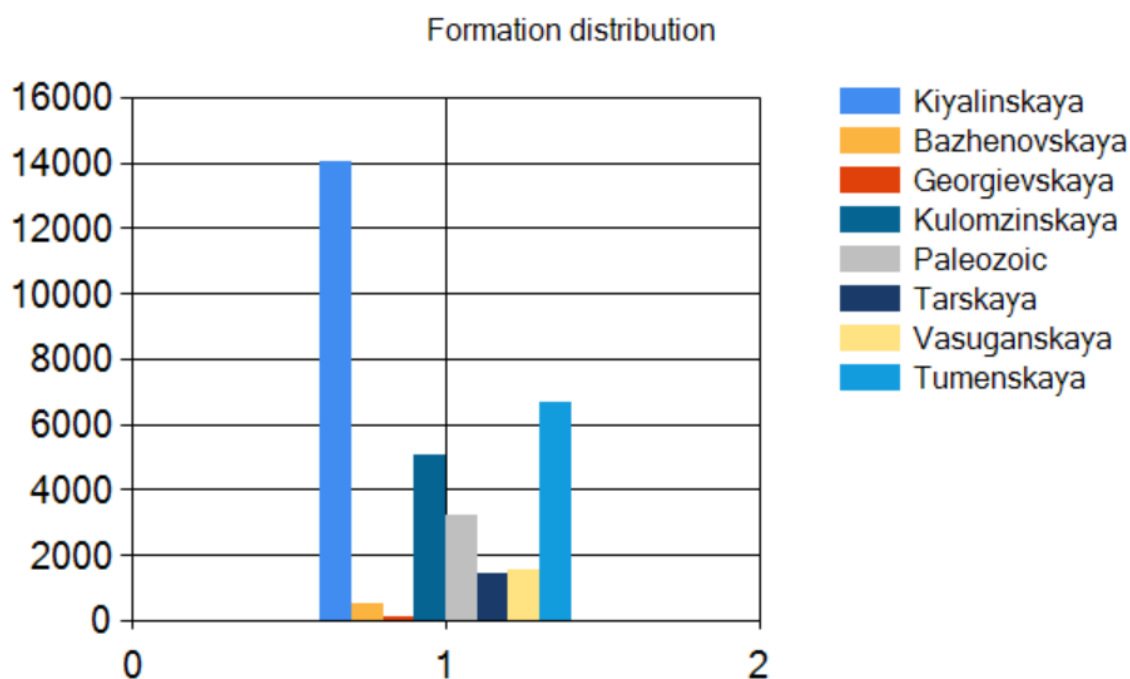


Figure A.4 – Histogram of the distribution of data on the suites

According to Figure A.4, one more drawback of the initial data is seen – different number of training vectors in classes. To solve this problem, cross-entropy is used as optimization criteria, which reduces of classifier uncertainty.

Logging measurements are given by points with a fixed depth step, but it can be different in different wells. When compiling a training sample, the log data must be sampled in absolute depth. The sampling rate is 0.1 m. The more detailed the sampling rate, the more detailed the geological profile will be.

When solving classification problems, three main types of tasks are distinguished:

- binary classification (positive and negative sign).
- multi-class classification (the object has attributes of only one class);
- multi-label classification (the object has the attributes of several classes simultaneously).

The task of selecting reference horizons refers to a binary classification, and the problem of identifying a suite in a well by log data is multi-class. For binary classification, the output vector is a scalar value: 1 - marker, 0 - not marker. In the problem of classification of log data for the suites, 8 classes were defined:

Kiyalenskaya, Bazhenovskaya, Georgievskaya, Kulomzinskaya, Tarskaya, Vasyuganskaya, Tyumenskaya, Paleozoic deposits. There are two ways to encode variables of a nominal type: Label Encoding and One Hot Encoding. For classification tasks, the best option is to select the One Hot Encoding method, in which the size of the output vector is equal to the number of classes, and the i -th component of the vector indicates the probability of having the class i of the object and takes a value in the range determined by the activation function of the output layer neurons.

An important issue is to determine what is the input vector, which will actually be fed to the classifier's input. A number of works are based on the point-by-point compilation of the training sample, i.e. The n -dimensional vector of attributes - values of n -different logs for a given depth of the well is fed to the input of the network. Thus, the more significant parameters are in the learning vector, the more accurately it will be trained.